



L. Randall Wray

UNDERSTANDING MODERN MONEY THEORY

Money and Credit in
Capitalist Economies

Understanding Modern Money Theory

Understanding Modern Money Theory

Money and Credit in Capitalist Economies

L. Randall Wray

*Senior Scholar and Professor of Economics, Levy Economics
Institute, Bard College, USA*

 **Edward Elgar**
PUBLISHING

Cheltenham, UK • Northampton, MA, USA

© L. Randall Wray 2025

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system or transmitted in any form or by any means, electronic, mechanical or photocopying, recording, or otherwise without the prior permission of the publisher.

Published by

Edward Elgar Publishing Limited
The Lypiatts
15 Lansdown Road
Cheltenham
Glos GL50 2JA
UK

Edward Elgar Publishing, Inc.
William Pratt House
9 Dewey Court
Northampton
Massachusetts 01060
USA

Authorised representative in the EU for GPSR queries only: Easy Access System
Europe – Mustamäe tee 50, 10621 Tallinn, Estonia, gpsr.requests@easproject.com

A catalogue record for this book is available from the British Library

Library of Congress Control Number: 2025933451

This book is available electronically in the **Elgaronline**
Economics subject collection
<https://doi.org/10.4337/9781800375154>

ISBN 978 1 80037 514 7 (cased)
ISBN 978 1 80037 515 4 (eBook)

Contents

<i>List of figures</i>	vi
<i>Preface</i>	vii
1 Orthodox versus heterodox approaches to money	1
2 Origins and nature of money	14
3 Credit money and the rise of banking, central banking, and public money	34
4 The state theory of money	66
5 Introduction to the origins of the endogenous money approach	94
6 Endogenous money: a credit and state money approach	122
7 Money and interest rates	148
8 Sovereign currency and fiscal policy	175
9 The balance sheet approach to money and finance	206
10 Money and the logic of capitalism	238
11 Employment policy and the value of the currency	265
12 Conclusion: the challenges ahead	292
<i>Bibliography</i>	304
<i>Index</i>	316

Figures

8.1	Growth and the federal budget balance, 1930–2022	193
8.2	Growth and the federal budget balance, 1998–2009, 2009–2020, 2020–2022	194
9.1	Balance sheets for each sector	208
9.2	Bank lending finances production, firm issues bond to retire loan	208
9.3	Commercial bank borrows reserves to meet requirements	210
9.4	Bank makes loan to finance investment	212
9.5	What happens to profits if there is no net investment	219
9.6	Balance sheets with a government sector	221
9.7	Impact of treasury spending	222
9.8	Treasury receives tax payments	223
9.9	Treasury sells bonds to central bank	224
9.10	Treasury sells bonds to banks	225

Preface

I began working on my first book manuscript under the tutelage of Jan Kregel while on a Fulbright fellowship in Bologna, Italy, in 1986–7. Edward Elgar had just formed his own publishing company and (apparently) asked Jan if he knew of any manuscripts that might be available. Jan recommended mine and Edward took a chance, publishing it in 1990 as *Money and Credit in Capitalist Economies: The Endogenous Money Approach*. The book explored in detail the history of the endogenous money approach, showing that the then-dominant exogenous money theory with its deposit multiplier and central bank control of the money supply was an aberration. The endogenous money approach had been commonly embraced throughout the 19th century, and I wanted to show that the domination of the exogenous money approach in the second half of the 20th century was a wrong turn.

In spite of Hyman Minsky's warning to me to avoid writing a *Genesis* story of money's origins, I did dip into speculation on the topic. While in Bologna, I met Otto Steiger, who was developing a private property approach to money and interest, identifying money's origins in Mesopotamian loans of grain. Jan directed me to Keynes's *Treatise on Money*, which led me to Knapp's *State Theory of Money* and finally to what Keynes called his period of *Babylonian Madness* when he discovered that the names of all the early money units came from measuring units for grain. As I was also familiar with Lewis Henry Morgan's (1877) studies of tribal society and the rise of class society, I was naturally skeptical of the orthodox story that claimed money originated as a medium of exchange created to replace inconvenient barter.

Still, the main topic of my first book concerned how money *works* in the modern capitalist economy. I also looked in detail at recent innovations in finance, including the dangers of the increasing use of *off-balance-sheet* commitments, the rise of securitization, and the growing competition from *non-bank* banks (what we now call *shadow* banks). I warned that while all these innovations might enable individual institutions to shift risk, they made the system as a whole riskier and more likely to collapse—two decades before the Global Financial Crisis when that came to pass.

In 1997, Edward Elgar took a much bigger risk, agreeing to publish my second book, *Understanding Modern Money: The Key to Full Employment and*

Price Stability (1998). By then, I had come to know Edward and Sandy Elgar, and their darling daughters, quite well, as we regularly caught up at the book displays at academic meetings. While working on the new book, I tried to convince Edward it would revolutionize the way we look at money—paraphrasing what John Maynard Keynes had said seven decades earlier about his *General Theory*: “I believe myself to be writing a book on economic theory which will largely revolutionize—not, I suppose, at once but in the course of the next ten years—the way the world thinks about economic problems.” Further, I asked him to price it at \$25 and guaranteed it would sell.

That book laid out what became known as *Modern Money Theory* (MMT). The term “modern money” was an inside joke, taking off from Keynes’s statement in the *Treatise* that Knapp’s *state money theory* applied to all money “for the past four thousand years”—i.e. the *modern* period.

Unfortunately, the reception was not entirely positive! Some heterodox economists accused me of abandoning the ideas advanced in the first book, insisting that the *state money approach* of MMT is inconsistent with the *credit theory of money* addressed by the endogenous money approach. The heterodox academic journals generally refused to publish scholarly work on MMT. Orthodox economists accused MMT of claiming that the government should print up gazillions of dollars, causing hyperinflation. MMT’s advocacy for using the *Job Guarantee* to ensure full employment was rejected as some form of forced labor. For a long time, policymakers and media pundits insisted that MMT was *crazy talk* and certain to destroy the country.

But when challenged to find holes in MMT’s explanation of *how the government really spends*, or why MMT claims that *a sovereign government cannot run out of its own currency*, all the critics had was *crickets*. No one has found any holes in MMT’s logic or description.

In the long and slow recovery from the Global Financial Crisis, MMT gained followers and proponents as the ideas spread across blogs and the internet more generally. They put pressure on politicians, asking why the government refused to do more to alleviate suffering. With ample unemployed resources, why couldn’t the government use its sovereign currency? Followers no longer believed President Obama when he claimed he’d like to do more, but the *government had run out of money*. They criticized prominent pundits like Paul Krugman for *dovish* hesitancy in supporting more spending.

The extremely deep COVID recession brought MMT into the spotlight again. After so many years of ridicule, suddenly politicians *embraced* MMT to justify spending trillions of dollars, euros, and yen to relieve suffering. While MMT proponents had recommended a different, targeted response, it was still good to see government spending for the first time since WWII without worrying about *pay-fors* (raising taxes to finance the spending). However, the accolades were relatively short-lived because while recovery was swift, inflation

rose. Suddenly, MMT was blamed for the inflation—that was largely due to supply chain disruptions and price gouging.

So, to some extent, the criticism is unfair—the policies adopted were not those recommended by MMT. On the other hand, it is also due to a failure to communicate the *theory* that underlies MMT. MMT has largely come to be seen as a *description* of government finances and a *policy* of turning on the spending tap to push the economy to full employment. MMT is said to claim that *deficits don't matter*, that *there's no crisis so deep that it cannot be solved by spending more*, and that *there's no reason to worry about inflation, because we'll just raise taxes*. Often, MMT is reduced to overly simplistic slogans: *taxes drive money*; *bond sales are just a reserve drain*; *the government spends the currency first so that people use it to pay their taxes*; and *the government sets the prices when it decides how much it will pay*. While there's an element of truth in all these statements, each comes with important caveats.

Sometimes the basic principles of MMT are said to be universally applicable—around the world today, and back through historical time. However, I was influenced by Minsky, who argued that any social theory that claimed to be universal would be useless. Economics is the study of an extremely complex social system—what we call *the economy*—that is embedded within an even more complex social system that includes political, cultural, and religious components. And the system is dynamic, continually evolving from a only partially understood past toward an unknowable future. For these reasons, useful social theory must be specific to the institutions that inform, shape, and guide human behavior.

Orthodox economic theory claims to be universal—it explains how an economy with *rational economic agents* would operate. As we'll see, it introduces money in a particular manner that is consistent with the universalist approach—beginning with an economy much like ours that has not yet discovered money and then adding money to lubricate the market mechanism. Sometimes MMT proponents have done the same thing—imagining that an authority imposes a tax that can only be paid in the authority's own currency to create the initial demand for money. As a logical exercise, it makes a point: those who owe the tax will work to obtain the currency needed to pay it. As a description of the historical origins of money, or of the role of money in the economic system today, it falls far short. It sheds little light on the actual creation and evolution of the monetary system we have today.

It is not sufficient to have unassailable logic plus a factually correct description of real-world operations. Understanding capitalism requires a theory of the way it works, and theories are formulated within paradigms. MMT must be situated within a paradigm—a worldview appropriate to studying the social system within which we live.

Before the pandemic hit, Elgar (the company) approached me about writing a second edition of the 1998 book. I suggested instead writing a new book that would integrate the ideas from *Money and Credit* with those of *Understanding Modern Money*—to examine both *credit money* and *state money* in a capitalist economy. Hence, this book.

My goal is to begin with the origins and nature of money—not to make statements of universal truth but instead to understand the role money plays in the dominant economic system today: *Capitalism*. When the Soviet Union fell, Minsky said it was finally OK to use the word capitalism; prior to that, only capitalism's *critics* used the term. However, Minsky used to joke that there are *57 varieties of capitalism* (adapted from the Heinz pickle slogan)! Still, these share common features that make capitalism quite distinctive and fundamentally different from previous kinds of economic systems.

What I will argue is that capitalism is a *monetary production system*. (See Dillard 1980.) This is not novel—it is the approach of the three great heterodox traditions that identify with Marx (Marxist, of course), Veblen (Institutionalist), and Keynes.

In the case of Keynes, we need to be specific because most of what passes for *Keynesian* is not consistent with a *monetary theory of production*—it is variously called *orthodox Keynesian*, *bastard Keynesian* (by Joan Robinson), *textbook Keynesian* (as it long dominated the textbooks), or the *Neoclassical Synthesis* (a term apparently coined by Paul Samuelson during the McCarthy period when Keynes was seen by Americans as a socialist). Instead, the followers of Keynes who do adopt his monetary theory of production are called *Post Keynesians*.

What is the monetary theory of production? It is based on the distinction between the dominant forms of production in all other economic systems versus the form production takes in capitalism. All human societies have engaged in behavior that we can call economic—they had to produce the necessities of life: at a minimum, food, clothing, and shelter. Of course, what is viewed as necessary differs among societies and changes over time. We must include, today, healthcare, recreation and entertainment, formal education, and adequate retirement.

It is not always easy to identify economic behavior—it is often embedded in other behavior, including, for example, keeping the gods happy (superstition and religion) so that they protect us. It is more transparent in a modern capitalist economy: I go to *work* at 8:00 am and end my workday at 5:00 pm—that is my contribution to production—and I go shopping after work, contributing to market sales. Then I cook dinner for my family, which is not considered economic behavior although it sustains life. There is a stark separation between participating in *the economy* and engaging in noneconomic behavior. A key distinction is whether the activity involves earning or spending money.

The monetary theory of production focuses on the critical decisions made that drive the economy. Marx put it in the most revealing way: production begins with money to make more money; Keynes insisted that producers have no other goal than to end up with more money than what they started with. Veblen argued that in capitalism the business enterprise is only interested in *pecuniary* gain. In short, production must begin with money, and to be successful, it must end up with more money—to cover all monetary expenses with money profits left over.

That leads us to the main questions that guide the analysis in this book: What is money? Where does the initial money that starts the production process come from? What determines whether *more money* will be generated by the production process? *Who* gets to keep the extra money?

Those are the questions that I've puzzled over for more than four decades. The two previous books tackled them, but not always in a straightforward manner. This book will present a more coherent argument that is also consistent with more recent scholarship. As indicated above, I have long wondered about the origins and nature of money, and my views have evolved. In my first book, I was focused on the origins and source of private credit money, while the second book concentrated on the state's money. In this book, I will address the interplay and joint evolution of the two that led to the *modern monetary systems* of the developed capitalist countries. I will also discuss the origins and evolution of the capitalist *system*. Just as our views of money needed to undergo a revolutionary change, our understanding of the origins of capitalism has undergone an eye-opening revolution in recent years.

In my 1990 book, I had a section that discussed the role of money and *the logic of capitalism*, a topic that I will take up in much greater detail in this book. Ultimately, this is the key feature of capitalism that makes it truly different. Money has been used for “four thousand years, at least,” as Keynes put it, and markets have existed as long as civilization. Capitalism is often referred to as a *market system*—but that cannot be its identifying feature because it shares that with all civilizations. I will argue that capitalism has a *logic* that is driven by the role that money plays in organizing the economy. This makes capitalism uniquely suited to study as an economic system—a social system in which the economy is disembedded to a degree we do not see in other social systems.

While most of the heterodox response to my first book was positive, Victoria Chick raised an objection that, while the title indicated it was about “money and credit,” she complained that I did not distinguish between the two. I think she was right. At the time, I was using the term *money* to refer to the medium of exchange that passes from hand to hand, and *credit* to refer to balance sheet entries—on which the debate about endogenous money was focused. I would not make that distinction anymore. There is no money without credit.

What is the *nature of money*? As Keynes insisted, we need to distinguish between money as a measuring unit and money itself. As a unit of measure, a dollar is like an inch—it measures money *value* while the inch measures *length*. Most early money measuring units were adapted from measuring quantities of grain—including the British *pound*. Units of measure must be socially recognized and can be changed. When the USA declared independence, it dropped the British pound in favor of the *dollar*. Ironically, the British dropped the *pound* as a measure of weight in favor of the *kilogram*, while the USA stubbornly insists on keeping the *pound*. Clearly, the choice of measuring units is a political/legal affair. The *money of account* is generally chosen by the authority—it is the state's money.

What does the money unit measure? Succinctly, but rather unhelpfully, we said it measures value—the money value. The money value of what? Keynes said it measures the money value of the money *thing*, the form in which the *record* of money value is kept. This can be a clay tablet, a notched tally stick, a stamped coin, a paper note, or, today, an electronic entry on a balance sheet. While more helpful, it doesn't seem to fully answer the question: *value of what?*

While we cannot be sure precisely how the grain measuring unit came to be used as a money measuring unit, the historical record suggests it was to keep track of debts of grain owed by Babylonian temples to workers. Throughout the history of money, what is being recorded is typically (but not always) an *obligation*, a debt. There are always two sides to a financial debt: there is a creditor and a debtor, so the debt is entered on the liability side of the debtor and on the asset side of the creditor. The exception to this is the value of a real asset, which is entered as an asset with a value entered in money terms on the balance sheet of the owner but appears on no balance sheet as a debt (to be clear, the money value is on the asset side and adds to net worth that is entered on the liability side of the owner).

When most people refer to money, they are thinking of the money record itself—the coin, the paper note, or the entry by the bank on one's demand deposit account. What all of these have in common is that they are debts (of the issuer) measured in money terms. This should be clear, but what does it have to do with value? Many would jump to the conclusion that the value of just about anything can be measured in money terms—not just debts. Everything has its price in terms of money. Is value just another word for *price*?

Discussion of value theory is usually a conversation-killer among economists. It is considered to be an old-fashioned notion put forward by Marx that even Marxists today have largely dropped as fundamentally flawed. The so-called *transformation problem* prevents us from mapping value to price; thus, value is rejected as unobservable and ultimately irrelevant. All we are left

with is money price—there is no external measure we can use except money. Money's value is simply the inverse of the price level.

This book will take a different view. To understand the internal logic of capitalism, we need an external measure of value for money itself. Indeed, we will argue that it is useful to employ two *theories of value* to understand that logic. We will use Marx's *labor theory of value* as well Keynes's *theory of the value of assets* (liquidity preference theory). It probably would be useful to add other concepts of value, depending on the question under analysis. For example, we could use the Institutional *instrumental value theory* if we were to go deeply into policy analysis, but that is beyond the scope of this book. We could also get into *moral values*—a topic economists usually shun; however, from Adam Smith to Keynes, many saw economics as a branch of *moral philosophy*.

We will stick with two—money value and labor hours—while leaving room for other value theories where they might help. Marx and Keynes both explicitly used those two measuring units in their analyses: *money* and *labor hours*. We will argue that for the purposes of analyzing an economic system based on monetary production, there really is no other choice of measuring units. Consistent use of these will enable us to reveal the logic of monetary production. Understanding that logic will help us to answer the questions that I've pondered over the decades: What is money and where does it come from? What determines whether *more money* will be generated by the production process, and who gets to keep the extra? What role does government play in the monetary production economy, and how can we improve policy to promote economic, financial, social, and environmental sustainability?

Those questions set the agenda for this book. They are the questions that naturally follow from the adoption of a monetary theory of production.

*

We've come a long way since 1988 when I was working on my first book with Edward Elgar Publishing. In addition to my 1998 book, I published a number of volumes with Elgar, among which my 2004 book, *Credit and State Theories of Money: the Contributions of A. Mitchell Innes* (also a book on the tradition that was incorporated into MMT) is my favorite. More recently, I was one of the editors of another great Elgar book, *Modern Monetary Theory: Key Insights, Leading Thinkers* (2023), and a co-editor (with Yeva Nersisyan) of *The Elgar Companion to Modern Money Theory* (2024)—essentially an encyclopedia of entries on topics related to MMT, written by the foremost authorities.

In addition to all the books Elgar has published on MMT and heterodox economics, the company has been a supporter of Post Keynesian and MMT

research and study over the years—providing funding and book displays at various events.

Meanwhile, over the past four decades, Elgar—the company—has become, arguably, the premier publisher of books on economics. I am grateful for the chance that Edward Elgar—the man—took with me.

I want to thank my mentors, including John Henry, Joe Furey, and Marc Tool at California State University, Sacramento, Hyman Minsky and Steve Fazzari at Washington University, and Jan Kregel, who helped supervise my work in Bologna. I benefited from the friendship and inspiration provided by Kenneth Boulding, Dudley Dillard, Gladys and Fagg Foster, Warren Samuels, John Adams, David Levine, Geoff Ingham, Duncan Foley, Charles Goodhart, Geoff Harcourt, Wynne Godley, Bill Dugger, Michael Hudson, Jamie Galbraith, Chris Desan, Rohan Grey, David Fox, and many others who have had a great influence on me.

I also thank my former students who have become close friends and collaborators, including especially Pavlina Tcherneva, Yeva Nersisyan, Eric Tymoigne, Flavia Dantas, Alla Semenova, Fadhel Kaboub, Stephanie Kelton, Zdravka Todorova, and Scott Fullwiler; former colleagues Peter Ho, Robert Urquhart, Tracy Mott, and Jim Sturgeon; fellow MMT travelers—especially Bill Mitchell, Mat Forstater, Martin Watts, and Warren Mosler; Maren Poitras for her tremendous MMT documentary (*Finding the Money*); those who have read drafts of the book's manuscripts or provided sources including David Freund, Eric Tymoigne, and Scott Fullwiler; and my Levy Institute colleagues—above all Dimitri Papadimitriou for all of his support over the years. I thank Eric Lin and Yeva Nersisyan for help with figures used in the book and Alan Sturmer for decades of editorial advice and monumental patience as I missed deadlines year after year.

Finally, I want to recognize the sacrifices made by my wives, Shona Kelly Wray and Xinhua Liu Wray, and children Shane, Alina, and Allison, that gave me time over the past 35 years for the research that went into this book.

This book is dedicated to Edward Elgar, the man who has been so important to my career and to the careers of many other budding economists.

1. Orthodox versus heterodox approaches to money

Economic textbooks often begin with the “economic way of thinking”—providing a framework for doing economic analysis. It starts with the individual who is maximizing utility and the firm that maximizes profits. Perfect competition is assumed—so many buyers, so many sellers that none can influence the price. Free markets with flexible wages, prices, and interest rates (and exchange rates when a foreign sector is introduced) seek market-clearing equilibrium. Later in the textbook, it is admitted that there could be “market failures”—largely due to rigidities (of wages, prices, interest rates, and exchange rates) but also due to imperfect information. Government is finally brought into the analysis about halfway through the text, as its intervention is both a source of those failures and a possible solution to some of them.

Our concern in this book is the nature and role of money in the economy we live in today, but orthodoxy’s treatment of money is schizophrenic. It is first introduced as a cost-reducing medium of exchange that mediates trade and should be neutral (affecting only nominal variables), but various devices are admitted into the analysis that can allow for temporary real effects. Inflation is created by government as it tries to take advantage of money’s transitory non-neutrality, but otherwise, we should not worry much about it. Indeed, the most advanced orthodox approaches to economic analysis at every stage of the evolution of its technique—from the original Neoclassical approach, to General Equilibrium Theory, through the ISLM “Keynesian” approach, on to various versions of Monetarism (Friedmanian, New Classical, and New Monetary Consensus) and to Real Business Cycle and Dynamic Stochastic General Equilibrium Theory—can find little reason to focus on money. According to orthodoxy, money just shouldn’t matter.

And yet, when we look at real-world debates outside academic economics, many of the most important disagreements have involved the role of money in the economy. In policy-making circles, money *certainly* matters. For example, the currency school-banking school controversies of the early 19th century, as well as Lincoln’s Greenback experiment, followed by free banking and William Jennings Bryan’s “cross of gold” campaign, were all about money. Money was blamed for the Great Depression when Roosevelt had to declare a

bank holiday to shutter all the banks and put Jessie Jones in charge of deciding which ones could reopen. The debates at the time claimed either that there was too much money (spurring the *Narrow Banking* proposal backed by Friedman, Simons, and Fisher) or too little (Friedman and Schwartz blamed the Fed for stupidly reducing the money supply)—but in either case, the depression was money’s fault. Something surely went wrong with money, again, in the early 1980s when Chairman Volcker promised to reduce its growth, resulting in interest rates above 20 percent, the failure of half of the US thrifts, and the developing country debt problems that led to the probable insolvency of the biggest US banks by the early 1990s. And then came the Global Financial Crisis of the mid-to-late 2000s, requiring a \$29 Trillion bailout by the US Fed. (Felkerson 2011; Tymoigne and Wray 2014)

MAYBE MONEY DOES MATTER—AFTER ALL?

In this introduction, we will quickly summarize the orthodox view of the origins and nature of money and then begin to outline the heterodox approach, highlighting money’s origins in debt. Money is traditionally linked to trade in markets; however, we will argue that the existence of money long predates market exchange. Still, it is important to emphasize that money plays a much more important role in the capitalist economic system (often erroneously referred to as a “market system”) than in any previous system. Capitalism is a monetary system; indeed, it is a *system of monetary production*. We cannot analyze capitalism as if it were a system of exchange to which money is added to lubricate the market mechanism. It is thus important to investigate the early beginnings of capitalism and the logic of monetary production economies. This will then provide the framework for analysis of *credit and state money* in the rest of this book—and to draw a sharp distinction between that framework and the approach of orthodoxy.

Throughout this book, we will contrast the monetary theory of production approach with the orthodox approach, emphasizing how differently money is treated. This addresses the main question of the book: what role does money play in a *capitalist* economy? The orthodox approach sheds no light. The heterodox tradition, following the lead of Marx, Veblen, and Keynes correctly situates money within a monetary production approach.

ORTHODOXY

The orthodox approach to money typically begins with an origins story based on the evolution from primitive exchange to the modern economy. Paul Samuelson, the writer of the dominant economics textbook of the early post-war period, nicely summarized the view:

Inconvenient as barter obviously is, it represents a great step forward from a state of self-sufficiency in which every man had to be a jack-of-all-trades and master of none. ... If we were to construct history along hypothetical, logical lines, we should naturally follow the age of barter by the age of commodity money. Historically, a great variety of commodities has served at one time or another as a medium of exchange: ...tobacco, leather and hides, furs, olive oil, beer or spirits, slaves or wives ... huge rocks and landmarks, and cigarette butts. The age of commodity money gives way to the age of paper money ... Finally, along with the age of paper money, there is the age of bank money, or bank checking deposits.

(Samuelson 1973, pp. 274–6)

Although this explanation of the origins of money and financial institutions is still taught in most money and banking courses, it is internally inconsistent and has no historical foundation.

According to the orthodox story told by Samuelson and others, barter replaced self-sufficiency and increased efficiency by allowing for specialization. One person could specialize in fishing while another could focus on farming; they then swap fish for corn—so long as the fisher wanted corn, and the farmer wanted fish. It was then discovered that further efficiency could be gained by using some object as a medium of exchange to eliminate the necessity of a happy *coincidence of wants* required for barter to take place. An evolutionary process would select the best-suited commodity to function as money—typically gold or silver—because of its physical characteristics (durable, high value to weight, scarcity). Further, *fairground barter* (where producers brought their wares to trade) replaced *isolated* barter because this reduced the time taken to complete a transaction. Thus, money springs forth to facilitate exchange by *lubricating the market mechanism*, and the creation of formal markets with higgling and haggling together reduces transaction costs, moving the economy toward its optimum position. (See Goodhart 1998.)

The argument is extended to the development of paper money by noting that in the 17th century, gold was commonly deposited with goldsmiths for safe-keeping against receipts called *goldsmiths' notes*. Time and effort (called *shoe leather costs*) could be saved by exchanging notes rather than by reclaiming the gold each time an exchange was desired. Goldsmiths realized they could safely lend gold or issue additional receipts as loans. However, as goldsmiths had to keep some commodity gold to facilitate clearing with other goldsmiths and for deposit withdrawals, the quantity of notes issued would be closely governed by the gold held in reserve. Some goldsmiths gradually specialized as bankers, and the modern private banking system emerged, based on fractional reserve deposit banking. Since the cost of issuing receipts was less than that of mining gold, goldsmith “banking” was another rational economic decision

taken to reduce the costs of the transactions structure; paper money thus largely replaced commodity money in circulation.

Government's role comes along later in the story—as it needed money to operate, it could impose taxes to be paid in gold, monopolize the gold mines, or claim the sole right to mint gold coins (in which case it could use seigniorage rights to obtain purchasing power by keeping a portion of the coins it minted on behalf of those with gold). Stamped coins reduced transaction costs so long as the government's stamp could be trusted as an indication of purity. The government might also follow the lead of the goldsmiths, issuing paper notes fully redeemable in gold. But the temptation to debase the coins or to reduce the gold backing notes would prove too strong to resist—leading to devaluation of the government's currency and inflation.

Eventually, this devolved to fiat money—with value determined by some combination of fiat (legal tender laws requiring acceptance of the government's currency) and trust—issued either through government treasuries or their central banks. Private banks were required to hold the fiat money as reserves against their own banknote issues. Thus, an increase in the issue of government fiat money would lead to a multiple expansion of banknotes in the fractional reserve system. Eventually, banknotes were replaced with bank deposits, but again, backed by reserve deposits held at the central bank.

While the deposit multiplier might vary, central bank control over the privately issued demand deposits is ensured through control of bank reserves. (See Brunner 1968.) To prevent excessive money from being privately created, the central bank must closely regulate the quantity of reserves. Lack of moral fiber on the part of the authorities leads to excessive reserves and money (both private and government). When the public finds itself with too much money, it spends the excess, causing inflation. Thus, the primary responsibility of the central bank is to prevent inflation through careful control of the money supply.

CRITIQUE

The orthodox story, in which the present is a linear descendant of the past, relies critically on deductive logic. The orthodox economist views our economy as a more-or-less free market economy in which only real variables matter (at least in the long run) and in which neutral money (affecting only nominal values) is used primarily to facilitate the exchange of real goods, undertaken by self-interested maximizers for personal gain. The origins of money are then discovered by abstracting from this hypothetical economy to one that is a close replication save one feature: it does not use money. The conventional economist compares these two economies and finds that the one using money faces lower transaction costs. Money must, therefore, have been created to reduce the transaction costs that arise in barter. Historical detail can be added to the

picture; mental gymnastics ensures that historical “facts” are consistent with the basic Neoclassical view of the world.

Money can be discovered in almost any society (past or present) if one is willing to include as money “tobacco, leather and hides, furs, olive oil, beer or spirits, slaves or wives ... huge rocks and landmarks, and cigarette butts,” as Samuelson does. If such objects cannot be found in a particular society, one can always argue that this society has not *yet* discovered money. All societies, according to this view, are based on exchange, or at least would be if natural propensities were allowed to flower. If one is willing to define almost any human interaction (even marriage—as Gary Becker 1974 argues) as an “exchange,” then exchange can be found in any society (even nonhuman ones). Finally, all such exchanges must be made on the basis of cold calculation of self-gain, for no other exchange would be rational.

Thus, the orthodox explanation of the origins of money is based on the existence of an economy based on barter exchange in formal markets (the fair-ground barter) which predates the introduction of money. But this is neither historically accurate nor is it coherent. The institutional prerequisites to the development of market exchange include the existence of private, alienable property, recognition of individual responsibility, self-interested behavior, at least some labor specialization, and forward-looking production. Yet, the historical examples of non-monetary exchange used to justify the traditional approach rarely show any of these characteristics.

Here we will argue first that primitive exchange or barter did not lead to the development of markets; second, that money did not develop out of primitive exchange; third, that both fiat money and credit money *predated* coined commodity (gold) money; and fourth, that the quantity of credit money has not typically been constrained by the quantity of central bank reserves, as in the multiplier story.

The orthodox stories of money’s origins are based on the supposed existence of primitive market exchange in early societies. Clearly, for a market exchange to occur, people must come to the market owning a diversity of items to trade—and there must be happy, self-interested, “double coincidences of wants” that lead to bargaining over relative prices (exchange rates)—at least for the type of market orthodoxy imagines. However, for most of their existence, humans lived in tribal societies where the exchanges that occur are “public acts performed in regard to the status of persons and other self-propelling things ...” (Polanyi 1971, p. 75); these exchanges have as their main aim to “exchange articles which are of no practical use ...” (Malinowski 1932, p. 860); sometimes “the identically same object is exchanged back and forth between the partners ... the sole purpose of the exchange is to draw relationships closer by strengthening the ties of reciprocity” (Polanyi 1971, p. 74); exchanges were frequently made to *equalize* wealth, rather than to achieve mutually beneficial

allocations of resources; and there was generally no fixed exchange rate among exchanged goods—the exchange rates would depend upon the status of the parties to the exchange. (Heinsohn and Steiger 1983)

And beyond these considerable barriers to the orthodox story, as Graeber argues, there is simply no anthropological evidence in support of neoclassical's version of barter in tribal society. There are instances in which barter is used as emergency behavior (so-called *cigarette money* in prisoner of war camps) or in trade *between* societies (often called *silent trade*). As Dalton notes,

moneyless market exchange was not an evolutionary stage in the sense of a dominant mode of transaction preceding the arrival of monetary means of market exchange. Barter occurs very widely in past and present economic systems, but always as minor, infrequent, or emergency transactions employed for special reasons by barterers who know of alternative and more important ways of transacting. (Dalton 1982, p. 188)

If these exchanges were not market exchanges, then what are the *primitive monies* (Samuelson's tobacco, huge rocks, and wives) that have been identified as the cost-reducing solution to barter exchange? These “primitive monies [are] used to create social relationships ... prevent group hostility and warfare ... elevate one's political position ... and restore peaceful social relationships between persons and groups disrupted by conflict ...”. (Dalton 1982, p. 185) These “monies” are always used in “special ways only” (ibid., p. 185) and never as a social unit of account; the “special purpose monies or highly ranked treasure items necessary to the transaction” may be used only in specific ways, and other items cannot be substituted for them in these specified transactions. (Dalton 1967, p. 264)

These “monies” did not function as media of commercial exchange, and they did not function as common measures of value, nor as the standard of deferred payment. Malinowski's study of the Trobrianders found:

It is obvious at once that in economic conditions such as obtain among the Trobrianders there can be no question of a standard of deferred payments, as payments are never deferred ... In fact, the narrow range of exchangeable articles and the inertia of custom leave no room for any free exchange, in which there would be a need for comparing a number of articles by means of a common measure. (Malinowski 1921, p. 13)

Similarly, while *loans* existed in nonmonetary tribal societies, these were fundamentally different from the forward contracts that characterize loans in monetary economies. First, in tribal society, *loans* are initiated by the *lender*, who forces the *debtor* to accept a gift. (Dalton 1967) Second, this loan is not undertaken with the prospect of material gain, for the chief motive is to obtain prestige by *destroying* one's wealth. (Heinsohn and Steiger 1989) Third, in

such societies, *repayment* terms are always fixed by social norms of reciprocity and redistribution—they are not the result of private negotiation. Thus, neither “money” nor “loan” represents economic phenomena (in the strict sense of the term)—rather, the purpose is identical to the purpose of *primitive exchange*: to reproduce social relations appropriate to tribal society through reciprocity and redistribution.

On the basis of historical and anthropological evidence, it can be concluded that primitive exchange was not market exchange *without money* and that the primitive monies did not emerge from the reduction of transaction costs in the exchange process. Dalton (1982) prefers to substitute the term *primitive valuables* for primitive monies in the analysis of most primitive societies since “demonstrably, most were not crude proxies for dollars or francs in simple market transactions.” (Dalton 1982, p. 183) While these may have served “as special means of commercial or noncommercial payment or exchange in primitive, peasant, and archaic societies all over the world,” (ibid., p. 183) they “are regarded as valuables to be used in special ways only; they are necessary means of reciprocal payment in social and political transactions.” (ibid., p. 185)

This does not mean that individuals in tribal society are completely lacking in self-interested behavior; rather, that such behavior would not normally be manifested in exchange for two reasons: first, since the community takes care of its members, gainful behavior in exchange is not necessary to provide a livelihood; second, reciprocity exerts continual pressure to eliminate self-interest from exchange *since it cannot benefit the narrow interests of an individual*. (Stanfield 1986, p. 59) Clearly, such exchanges do not conform to the orthodox view of profit-seeking market behavior but represent *conventional* (meaning following convention) behavior more akin to the Western practice of gift-giving at Christmas. (Bell and Henry 2001)

MONEY AND THE NATURE OF MARKETS

The orthodox story also suffers from a logical problem. Levine (1983) argues that the existence of a market requires diversity and variability of needs—otherwise, one must suppose that the existing distribution of endowments (resources, ability, know-how) just happens to be such that a significant portion of the population cannot produce for its own needs, so must produce for the market. The hypothesized barter economy assumes that individuals “organize their activities with the idea of marketing in mind” before money exists. (Levine 1983, p. 21) They specialize in producing commodities that they do not need to exchange in the market for desired commodities. But such production requires the pre-existence of private property and independence of individuals.

Furthermore, even production for (hypothesized) barter exchange in a private property economy involves time and uncertainty, but these are the additional prerequisites for *monetary* production—as discussed later. Thus, the prerequisites for a monetary economy already exist in the market economy that is supposed to be based on barter. Is it conceivable that barter could have predated the use of money, even though the hypothesized barter economy displays the conditions of a monetary economy?

Even if this can be accepted, it does not seem to generate the conditions required for an innovative, dynamic market, for production would be geared to satisfy the “historically developed and given modes of consumption and structures of neediness.” (Levine *ibid.*, p. 22) That is, the drive to accumulate is seriously constrained in the hypothesized barter economy. Only in an economy in which wealth is denominated in money does “need” become socially determined by the requirements of social accumulation.

Orthodoxy has not been able to explain how individual utility maximizers settled on a single numeraire. (Gardiner 2004; Ingham 2004a) While use of a single unit of account results in efficiencies, it is not clear what evolutionary processes would have generated the numeraire. According to the conventional story, the higgling and haggling of the market is supposed to produce the equilibrium vector of relative prices, all of which can be denominated in the single numeraire. However, this presupposes a fairly high degree of specialization of labor and/or resource ownership—but this pre-market specialization, itself, is hard to explain within the constraints of the conventional approach. (Bell, Henry, and Wray 2004) Once markets are reasonably well-developed, specialization increases welfare; however, without well-developed markets, specialization is exceedingly risky, so diversification of skills and resources would be prudent. It seems exceedingly unlikely that either markets or a money of account could have evolved out of individual utility-maximizing behavior.

Levine argues that in a monetary economy, individuals produce commodities they do not need because their motivation for production is the acquisition of wealth as such (i.e., making *money*, the pure form of wealth). The separation of producer from consumer supports a system of exchange. That separation comes about because the idea of wealth in general has become a part of a way of life and mode of thinking. The penetration of the idea of wealth into consciousness makes the pursuit of *more money* through commodity production an intelligible goal. (Levine 1983, pp. 22–3) This separation of producer from consumer requires (for its full development) the existence of private property and the creation of a class of propertyless workers—who must sell or rent (see Chapter 10) their time to obtain money. The existence of propertyless workers would extend market demand and the medium of exchange function of money. In conclusion, market supply is broadened as production becomes oriented to the market to obtain money-denominated assets; market demand

exists because propertyless workers must purchase the means of subsistence using money wages.

This is why heterodoxy rejects the claim that markets could pre-date money. Let's turn to the alternative.

THE HETERODOX ALTERNATIVE

The preceding discussion—of our critiques of the orthodox story based on the evidence—has admittedly flown through several thousand years, covering the whole period from the first tentative movements away from tribal society to the development of a capitalist society. On one hand, this can be justified by the argument that all pre-capitalist societies are more like one another than any is to capitalism. The origins of markets based on the use of money may be found in the early development of private property; however, money and production remained “embedded” in noneconomic social relations until the emergence of a “monetary economy” relatively recently.

There is no novelty in this claim; one can find similar arguments in Stanfield (1986), Polanyi (1957), and Heilbroner (1985). Stanfield argues that the major transition to the modern economy occurred in the 17th century, during which the Mercantilist state took an active role in the creation of internal markets; from this point forward, the economy gradually emerged and began to shape the rules and norms of society. (Stanfield 1986, p. 102)¹ This culminated, according to Polanyi, in the conscious creation of a “self-regulating free market economy” in the 19th century, in which the largely disembedded economy supposedly could function without the direction of authorities. A market myth was created, summarized as the belief that the pursuit of self-gain would achieve social provisioning even without purposeful pursuit of this social purpose. According to Polanyi (1957), the attempt at creating a self-regulating market economy failed, engendering a protective response, creating institutions and limiting the scope of markets precisely because they could not accomplish desired social provisioning.

Finally, Heilbroner (1985) argues that the creation of capitalist society represented a *revolutionary* movement in which an economic system is created whose overriding function is to accumulate *capital*, rather than to ensure social provisioning. This continual “expansive metamorphosis of capital” is the essential *logic* of capitalism. (op. cit. p. 36) Furthermore, this logic of accumulation takes the form of the accumulation of greater *nominal* values. Clearly, capitalism—a system based on nominal accumulation—is a system very different from previous institutionalized interactions among humans and between humans and nature. In capitalism, the overriding purpose of production is money; provisioning is secondary. (See Dillard 1980.)

However, one cannot ignore the substantial differences between tribal society—which does not use money—and pre-capitalist societies that do. Once private property and money appear, we have the development of markets—markets have been fairly common since the late Stone Age. (Stanfield 1986, p. 97) Monetary production, however, remains a much later development. As recognized by Polanyi (1971), markets exist long before a *system* based on markets appears. It is clear that markets can exist in a variety of economic systems—from the tribute economies based on central authority to the democratic societies of ancient Greece, and from the Roman slave society to the feudal societies of Europe. Even though most production in these societies never entered markets, markets and the sale of goods (and slaves) for money-denominated assets played a role. This admission, however, concedes nothing to the Neoclassical insistence on seeing a natural propensity to truck and barter in all societies. Rather, in most of these cases, markets were set up specifically by state intervention; in all cases, states regularly intervened.

Indeed, Levi-Strauss defines exchange as a peacefully resolved war and likens exchange rates to peace treaties. (Stanfield 1986, p. 90) Polanyi writes of the creation of markets by generals to provision armies; prices were not set by the higgling and haggling of markets. Rather, traditionally, trade carried no taint of commerce.

In its origins a semi-warlike occupation, it never cut loose from government associations, apart from which but little trading could take place under archaic conditions. ... Treaty prices were matters of negotiation, with much diplomatic higgling-haggling to precede them. Once a treaty was established, bargaining was at an end. For treaty meant a set price at which trading took its course. As there was no trade without treaty, so the existence of treaty precluded the practices of the market.” (Polanyi 1971, pp. 86–7)

The existence of private property, money, and markets is apparently not sufficient for the development of a market *system*, or, better, a system based on production for sale in markets for *money*. The extent of the market would remain limited until a substantial portion of the population became propertyless, thus becoming wage workers. As Stanfield (1986, p. 45) argues, the hungry must obtain food, but how they go about obtaining it is institutionalized. The propertyless, and therefore hungry, individual may rely on family, on begging, on government redistribution, or on wages by selling labor power (Polanyi’s fictitious commodification of labor), each depending on the response engendered by institutional arrangements. Bloch (1953) argues that the creation of a pool of potential wage laborers was accomplished by a process often (somewhat narrowly) called the enclosure movement. Again, however, the *labor market* did not spring naturally from this development, but was gradually created, and the extent of production for the market would be severely constrained until labor

itself became *commodified*—a process that resulted in part from accelerated privatization of land.

In the next chapter, we are going to argue that money emerged approximately 5,000 years ago, and not out of barter but more likely with the rise of civilization and a particular kind of debt. Money *preceded* markets, and the organization of the economy around markets—and monetary production—is much more recent. While the development of capitalism is often dated to the period between the thirteenth and seventeenth centuries (in Northern Europe; see Heilbroner and Thurow 1982), we will make the argument that capitalism as we know it is even more recent and that its most important institutions originated in the New World. Finally, the rise of capitalism is closely tied to the rise of the modern state—rather than seeing the state as an intervener into a well-functioning *free market* system, we will argue that the state supplies many of the most important institutions required for capitalism—above all, the state's money. This has been the primary focus of the Modern Money Theory analysis.

CONCLUSION

We have critiqued the orthodox view of money's origins and briefly introduced an alternative view of the origins of money, markets, and capitalism. This might seem tangential to our topic—the monetary production economy. However, an understanding of this alternative approach to history is important for revising our view of money's origins and nature, and the role that it plays in our economy.

The barter story—as well as Adam Smith's reference to a human propensity to “truck and barter”—makes it seem that production for markets with relative prices determined by “higgling and haggling” is natural. Selecting a commodity with particular characteristics (gold, the noble metal) to serve as a numeraire would reduce transaction costs, lubricating the market mechanism. The ability to hold gold for safekeeping would allow the goldsmith to naturally become the banker. A gold standard would keep money's value stable. The fractional reserve system with a deposit multiplier would support more economic activity, given the scarcity of gold. Eventually, sovereigns would intervene—dictating that their fiat money should become the banking reserve. This opens the economy to the dangers of inflation, excessive public debt, and potential insolvency. All of this is easily derived from the initial assumptions that lie behind the barter story.

However, this view is neither historically correct nor logically coherent. In the chapters that follow, we will dig deeper into our alternative view of money's nature, origin, and evolution. The alternative will be consistent with the historical evidence—such as it exists—and with our understanding of the way

the capitalist system operates. Most importantly, our economic system is fundamentally a *monetary* system in ways that will be discussed. While money has long existed in previous economic systems, those were not monetary systems as defined within the three great heterodox traditions: Institutionalist, Keynesian, and Marxist. Money plays a more fundamental role in the capitalist economy, and it makes capitalism quite different from all previous economic systems.

In the next chapter, we will argue that money's origins cannot be found in the dubious story of pre-existing market exchange based on barter. We will focus on three alternatives—all of which locate money's origins in debt relations. One posits private loans that eventually are denominated in a money of account. Another emphasizes tribal Wergild debts that are transformed into debts to authorities. The third argues that the money of account was created for internal record-keeping in Mesopotamian temples. In any event, the origins of money are uncertain—*lost in the mists of time*, as Keynes put it—but the existing evidence is that record-keeping in a money of account is at least *4000 years old* (again, as Keynes put it) and predates markets. Further, while markets have existed for thousands of years, the kind of *monetized* economic system we live in today—with monetized goods, services, assets, and labor—is of much more recent origin.

Chapter 3 examines credit money and the rise of banking, as well as the role of the modern state in the development of the banking system. Chapter 4 looks at the development of the Chartalist or State Theory of money and its links to the credit money approach—an integration embraced by MMT. Chapter 5 compares the orthodox *exogenous* approach to money with the early *endogenous* money approach of the Banking School, Marx, and Keynes, while Chapter 6 examines the more recent revival of the endogenous money approach by heterodox economists. This is also the approach embraced by MMT. Chapter 7 examines interest rate theory that is consistent with both the endogenous money approach and with Keynes's *liquidity preference theory*, distinguishing between the impacts of flows and stocks on interest rates. Chapter 8 provides a detailed look at the modern state money approach to sovereign currency and fiscal policy advanced by MMT. Chapter 9 presents a balance sheet approach to money—showing how banks create money, and how government spending is financed by coordination among the treasury, central bank, and private banks. Chapter 10 returns to an examination of the role that money plays in the capitalist system. It also integrates the labor theory of value associated most closely with Marx and the liquidity preference theory of asset prices that Keynes advanced. It argues that putting *value theory* back into our analysis of capitalism allows us to better understand its nature. Chapter 11 provides a detailed look at MMT's best-known policy proposal—the *Job Guarantee* program—and argues that it can act as a strong economic

stabilizer. The argument is linked to the previous chapter's approach to the determination of the value of money, discussing how the program can help to anchor the value of money. Chapter 12 concludes the book by examining the challenges we face.

NOTE

1. We will question the view that the origins of capitalism can be found in the dissolution of European feudalism—a notion that has long dominated economic thought. Instead, we will look at the origins of capitalism in the New World's slavery. See Chapter 10.

2. Origins and nature of money

If money did not originate as a cost-minimizing alternative to barter, what were its origins? In the following discussion, we will summarize research into the origins and early development of money. This is, of course, a difficult task. As the famous numismatologist Grierson notes, “Study of the origins of money must rely heavily on inferences from early language, literature, and law, but will also take account of evidence regarding the use of ‘primitive’ money in modern non-western societies. Such evidence, of course, has to be used with care.” (Grierson 1977, p. 12) He also recognized that the history of money is much more complex than the history of coins, for there is the danger that one might try to find money in societies that did not even use it. “Some systems, while employing shells or other commodities frequently used as ‘money’, may not necessarily be monetary at all!” (ibid., p. 13)

It is difficult for economists to agree even on a definition for money, and most recognize several different functions of money. It is possible that one might find a different *history of money* depending on the function that one identifies as the most important characteristic of money. While many economists (and historians and anthropologists) would prefer to trace the evolution of money used as a medium of exchange, our primary interest is in the origins of money of account, because, as it turns out, that is where money’s origins probably lie. This is essential to understanding money’s contemporary functions and its role in our economy. The evidence shows that the earliest money was not created to facilitate exchange but as the unit of account in which accounts were kept and debts were measured. For that reason, in this section we will speculate on the origins of money, and specifically, on the money of account.

Keynes argued in his studies on ancient currencies that “for most important social and economic purposes what matters is the money of account; for it is the money of account which is the subject of contract.” (Keynes 1982, p. 252) In the *Treatise on Money*, he argued:

A money of account comes into existence along with debts, which are contracts for deferred payment, and price lists, which are offers of contracts for sale or purchase. Such debts and price lists, whether they are recorded by word of mouth or by book entry on baked bricks or paper documents, can only be expressed in terms of a money of account. Money itself, namely that by delivery of which debt contracts and price contracts are discharged, and in the shape of which a store of general

purchasing power is held, derives its character from its relationship to the money of account, since the debts and prices must first have been expressed in terms of the latter ... Money proper in the full sense of the term can only exist in relation to a money of account. (Keynes 1971b, p. 3)

To be sure, we do not have decisive evidence for the precise origins of money as a unit of account, even if we can decisively reject the barter story. There are, however, several alternative explanations worth considering: one that locates money's origins in the rise of private property, one that emphasizes origins in the tribal practice of *Wergild*, and a third that focuses on internal accounting within Mesopotamian temples. All of these insist that money originated before and therefore outside market exchange and link its creation to the transition from tribal society to class society.

HETERODOX APPROACHES TO THE ORIGINS OF MONEY

We first go through the three alternative explanations for the origins of money, and then we will examine what light this sheds on the nature of money today.

1. The Private Property Story of Money's Origins

Let us begin with the view of Heinsohn and Steiger (1989) who emphasize the role played by private property. They argue that money and interest were developed at the same time—not out of a barter economy, but when private property and private debt were developed. The development of private property destroys the collective security of tribal or even command society and makes each member of society responsible for their own security. Private ownership makes loans from individuals possible, and individual responsibility induces creditors to require interest. (This links money and interest at the very beginning of the evolution of money—a point to which we will return later when we investigate Keynes's theory of interest and the monetary economy.)

Each proprietor is responsible to “set up his own reserves as a source of security which will protect him, in an unknowable future, from the need to ask credit of other—similarly isolated—proprietors,” which generates the drive to accumulate credits. (Heinsohn and Steiger 1989, p. 192)

With the establishment of private property, we at once have the elements of a money economy, that is an economy which is based on interest-bearing contracts between creditor and debtor and brings about monetary production. This production arises from the structural insecurity of the proprietors ... Private ownership production is a monetary production because the private reserves embodied in the creditor-debtor agreement are money of account. ... (ibid. 1989, p. 193)

When the private proprietor extends credit to others, he must be concerned with the risk to which he is now subject since he is giving up some of his security. Thus, a 'premium' must be charged, whose size is not determined so much by the volume of his reserves, but rather by his estimate of his existential uncertainty: i.e., by his estimate of the existential risk he runs during the period for which he will part with liquidity. (Heinsohn and Steiger 1989, pp. 192–3)

Money thus first appears in private property economies. These have occurred at various times and places in human history since the Greek city states.

Keynes shows that early monetary units were based on a specific number of grains of wheat or barley. (Keynes 1982, pp. 233–36) Thus, according to the interpretation by Heinsohn and Steiger, money was not created to facilitate exchange but rather as a debt contract after the development of private property. The first money existed when private property—say, a bushel of barley—was loaned with the expectation of payment of a greater sum in the future. Eventually, the measure used for barley would be used as a universal equivalent to act as the unit of account in all creditor-debtor relations to reduce transaction costs. Thus, the loan of a cow could be repaid in barley equivalents. Later, representative money, in the form of metal ingots, could be loaned, used as the terms in which debts were written, and used to settle debts. However,

This optimising of the storage potential of private property issued as credit, which eventually produced coinage does not represent an "essential", but merely technical element of the monetary economy ... (Heinsohn and Steiger 1989, p. 193)

Use of metal (or later, paper and tally sticks) as the medium for recording monetary debts depends on technology that changes over time.

The market, "then, is not a place of barter ... but a place for earning the means of settling debts, i.e. money." (Heinsohn and Steiger 1989, p. 195) Production occurs to retire debt, or to accumulate reserves of wealth to reduce the probability of becoming indebted. Money "is private property of lowest carrying costs which must be acquired by all members of the society in question to maintain or increase their (individual) security or to avoid falling too deeply into debt" (ibid., p. 196) A hoard increases individual security, whether that hoard is in terms of barley or in terms of access to another's reserves that is given by holding an IOU. An increase in concern regarding individual security raises the premium required to induce one to give up reserves. (All of this is similar to, or derived from, Keynes's theory of money and interest, as discussed below in Chapter 7.)

With the fall of economies based on private property and the rise of feudalism, the need for a money of account was diminished. Banks disappeared around the 3rd century AD, "and the medieval feudal command economy with its internally money-less estate began its more than thousand-year domination in

Europe.” (Heinsohn and Steiger 1984, p. 57) Even though money and exchange existed in feudal society (for example, in the medieval “fairs” to be discussed below), a monetary economy did not re-emerge until feudal society began to disintegrate in the late 14th century when private property reappeared. Money did not completely disappear as it was still used in some relations (particularly in foreign trade), but most production was for own use and responsibilities were shared by the communal feudal organizations. As feudal rents were gradually converted to money form, and as peasants demanded and gained rights to own land, the feudal system broke down—making way for the rise of capitalism.

This stands in stark contrast to the orthodox story in which markets based on barter arise naturally out of the utility-maximizing behavior of individuals, and in which money is created to facilitate exchange. In the Heinsohn and Steiger view, money was created in creditor-debtor relations (that require the existence of private property), but markets developed only with great difficulty and usually only with government support. With the collapse of the institution of private property after the fall of Rome, money and markets nearly disappeared.

According to Heinsohn and Steiger, money re-emerged with the development of private property and a class of propertyless workers during the evolution of capitalism. (Note, however, the discussion in Chapter 10 of the role played by slavery in the development of capitalism.) This not only leads to the development of a pool of employable workers but also leads to the possibility of the development of a market on a large scale. To produce for the market, the capitalist hires from the pool of propertyless workers. As production and sales take time, the capitalist must advance the means of subsistence, plus raw materials, to the workers. This requires either that they advance their own accumulated reserves (of food and raw materials) or borrow from those who are willing to relinquish reserves. Because the future is uncertain and failure is possible, borrowing reserves necessitates repayment of a greater amount to compensate the lender for the security relinquished. Thus, production for the market involves a “wheat now for more wheat later” proposition. Money as a unit of account naturally evolves out of this process, whether money as a medium of exchange takes the form of grain or coins. Given the uncertainty faced by independent property owners, interest must be paid on loans. (As we will see later, this is very close to Keynes’s view of the relations between money and interest.)

To summarize this position, money is a unit of account, created when private property arises, and an individual becomes a creditor or potential creditor. In the absence of private property, money as a unit of account is unnecessary because the community (or at least a portion of the community) is responsible for the well-being of individuals. Exchange may occur (e.g. the ceremonial exchanges of tribes), and payments may occur (e.g. *bride price* paid to the

wife's family for the loss of a woman to marriage, or the payment of a portion of the peasant's produce to the lord), but these societies have little need for money.

Although commodities with some of the characteristics of money did exist in tribal society, these differed in essential aspects from money. Most importantly, there was no fixed exchange rate between this *commodity money* (e.g., beads, shells) and other goods. Second, in societies without private property, when goods were loaned, often the society as a whole suffered the loss if they could not be returned. Individuals were normally not individually responsible for such debts and if return of a loaned item would result in individual hardship, its return was not expected. Third, in many instances, exchange occurred for the purpose of equalizing wealth, not to achieve mutually beneficial allocations of resources.

2. Origins of Money in Babylonia

In his analysis of ancient currencies, Keynes argued that even as early as the third millennium BC, one finds "very advanced indeed" the Babylonian use of money. He examined in detail the monetary reforms of Solon (circa 590 BC) and Pheidon (7th century BC) that set the values of coins. However, these values were based on weight units that could be traced back to 3000 BC, if not earlier. Keynes noted that the

mina, or mina, which Dungi prescribed for Ur in the middle of the third millennium BC is, within the limits of our positive knowledge, the earliest standard of weight. Recent discoveries have, however, thrown back the genesis of organised economic life to a date so much earlier than was previously supposed, that weights must have existed centuries, and, perhaps, even millennia before Dungi, in whose reign money, interest, contracts, receipts, and even bills of exchange are fully established ... (Keynes 1982, p. 232)

Indeed, Keynes argued that "the fundamental weight standards of Western civilization have never been altered from the earliest beginnings up to the introduction of the metric system" (op. cit., p. 239); without exception, "All weight standards of the ancient and also of the medieval world in Babylonia, the Mediterranean Basin and Europe have been based on either the wheat grain or the barley grain as their monad." (ibid.) The basic *monad* was then "60 x 60 x 3" grains of wheat for the *mina* of the "Egyptian system", or 60 x 60 x 2 barley grains for the "Lydian or Euboic system." (op. cit., p. 236) "Similarly, the avoirdupois grain is by contemporary definition the medieval wheat grain and the troy grain is the medieval barley grain." (op. cit., p. 237) Whether we speak of the *mina* and *shekel* or *pound* and *lira* of Europe, all the early money units were weight units based on either wheat or barley grains, with the nominal

value of gold usually measured in wheat units, and the nominal value of silver usually measured in barley units.

Monetary units, then, appear to be derived from weight units but do not derive their value from precious metal. Why weight units? It is possible that the weight units were just taken over because they offered well-known and objective standards. However, we know, for example, that “there is plenty of evidence for corn-wages and corn-rents from the Babylonian age onwards” (Keynes 1982, p. 258; Keynes is using the term *corn* in the generic sense to apply to grain), and for barley taxes in Mesopotamia (Hudson 2004). Is it possible that the choice of the wheat and barley grains as the bases of monetary units had a more concrete origin? And did they arise out of barter exchange or out of early debt relations?

Fortunately, we can find in Mesopotamia a nearly 10,000-year unbroken record of the development of both writing and money—indeed, the history of the two is closely linked. Writing seems to have evolved out of a system of counting, as Denise Schmandt-Besserat (2014) explains:

The Mesopotamian cuneiform script can be traced furthest back into prehistory to an eighth millennium BC counting system using clay tokens of multiple shapes. The development from tokens to script reveals that writing emerged from counting and accounting. Writing was used exclusively for accounting until the third millennium BC, when the Sumerian concern for the afterlife paved the way to literature by using writing for funerary inscriptions. The evolution from tokens to script also documents a steady progression in abstracting data, from one-to-one correspondence with three-dimensional tangible tokens, to two-dimensional pictures, the invention of abstract numbers and phonetic syllabic signs and finally, in the second millennium BC, the ultimate abstraction of sound and meaning with the representation of phonemes by the letters of the alphabet.

To summarize and simplify: writing was invented by accountants to keep track of debts. While the accounting was originally kept as tokens that looked like the thing counted (a token for a goat, a different token for a basket of barley), an early innovation was to simply push the token into fresh clay—the earliest form of *writing*. The next innovation was to use a stylus to carve a representation into fresh clay (dispensing with the tokens). Gradually, the writing became more stylized until the recording evolved into the *cuneiform* tablets we are more familiar with, and—importantly!—that we can read. Those earliest tablets were records of credits and debts. From those earliest examples of writing, it took about four thousand years to develop literary texts, and another thousand years to create an alphabet that allowed written language to become more like spoken language.

The conceptual leap required in writing was tremendous—and maybe the most important innovation of the past 10,000 years. Similarly, the conceptual

leap from recording debts taking the form of particular things owed (a goat, a basket of barley) to measuring debts in money values had to have been comparably difficult. While measuring units for length, volume, and weight were relatively easy to develop (so long as there was agreement to standardize them), the money-measuring unit was much harder. The record preserved in the clay tablets proves that this had been accomplished, at least in Mesopotamia, about five thousand years ago. (Not coincidentally, Graeber's famous book on money and debt is titled *Debt: The First 5000 years*. (Graeber 2011))

Michael Hudson (forthcoming; see also Hudson 2004) locates money's origins in the temple economies of Babylonia, created initially for internal accounting within the Bronze Age (3500–1200 BC) temples and palace communities. He describes how authorities developed cost accounting to coordinate the flow of food and raw materials needed to provision labor for public building projects and handicraft workshops. That provisioning required designing weights and measures. Prices were assigned monetary values in silver or grain to keep accounts and denominate fiscal balances. Silver and grain so valued were used for payments to the temples and palaces. These were the first forms of money used as a medium of exchange denominated in a money of account, whose origin is thus in institutional fiscal arrangements by authorities. According to Hudson, weighed pieces of silver were used for thousands of years in the Near East, usually in standardized weights often imprinted with the marks of the temples that produced the silver of specified purity (typically 7/8 pure). (Hudson forthcoming, pp. 9–10) However, these economies operated mostly on the basis of credit, rather than by paying in metal or barley:

Because of the time gap between planting and harvesting, few monetary payments were made at the time of purchase. When Babylonians went to the public alehouse, they did not pay by carrying grain around in their pockets. They ran up a tab to be settled at harvest time on the threshing floor. The ale women who ran these “pubs” then paid most of this grain to the palace for the consignments that had been advanced to them during the crop year. (Hudson forthcoming, p. 19)

In other words, they would pay down their money debts by delivering the appropriate quantity of grain equal to the debt, given the established price of grain. To provide a standard of value and serve as the means of payment, grain and silver had to be measured or weighed in standardized units. To facilitate calculation for resource allocation within the large institutions, these units were based on the administrative calendar created to allocate resources on a regular monthly basis. (Hudson forthcoming, p. 22)

According to Hudson, trade among merchants took place in the silver unit of account, while the grain unit “was designated as a unit of account to calculate values and co-measure labor time and land yields for resource allocation, and

also the means of payment for taxes and fees in the agricultural and handicraft sphere.” (forthcoming, p. 20) As Hudson puts it:

To value disparate commodities and facilitate these functions for grain and silver, a bimonetary system was created by setting the *shekel*-weight of silver (8 grams) as equal to a “gur” quart of grain or 300 *sila*. The acceptability of grain and silver for settling taxes, rents, and other official debt balances catalyzed their use as money throughout the economy. (ibid.)

Official prices, wages, interest rates, and taxes were set by the temple authorities, using round numbers to facilitate calculations. Months were standardized as 30 days (rather than the lunar month of 28 days), and the typical monthly allocation of grain to a temple worker was divided into 60 units, apparently to account for two meals a day. The sexagesimal system enabled the large institutions to calculate the rations needed to produce textiles or bricks, build public structures, or dig canals during any given period of time. Weights for silver and other metals followed suit by dividing the *mina* into 60 *shekels*. (forthcoming, p. 22)

The rate of interest on commercial advances denominated in silver was set in the simplest sexagesimal way: 1/60th per month (the decimalized equivalent being 20 percent per year), doubling the principal in five years (60 months). This standardized rate was adopted by the economy at large. (op. cit., p. 22) Contracts were recorded on *cuneiform* tablets. Some of these were encased with only a summary of the contents recorded on the case. When the contracted payment was made, the case could be broken to ensure that the safely ensconced tablet matched the record on the case.

In sum, the money of account was developed by temple authorities to value labor, output, and shares to be distributed. Prices in money were set by authorities and posted at the temples. Grain or weighed silver could be used for payment at the money value also established by authorities. The contrast with the barter story of money’s origins and with the supposed market determination of price should be clear. Money was created for internal accounting within the temples, and the value of money—prices in money—was established by the temples, not by markets. Rather, as markets developed, prices would be set by the temple in its money of account. (See Henry 2004 for a similar account of the origins of money in Egypt.)

Let’s turn to an intriguing third hypothesis on money’s origins: *wergild* fines. This was proposed by Grierson, the foremost scholar of coins.

3. Wergild Fines and the Origins of Money

The monetary measuring units may have first developed out of the elaborate rules governing *wergild*. Derived from “*wair*” meaning “man” and “*gildan*” meaning “to pay” or “to render” (Grierson 1977, p. 22), the term *wergild* refers to the ancient tradition of paying in-kind compensations for personal injuries to, or for murders of, kin members. “The general object of these laws was simple, that of the provision of a tariff of compensations which in any circumstances their compilers liked to envisage would prevent resort to the blood feud.” (op. cit., p. 19) By compensating the family of an injured or murdered person, retaliation by force and other “inconvenient social consequences” could be avoided. (ibid.) “Compensation in the Welsh laws is reckoned primarily in cattle and in the Irish ones in cattle or bondmaids (*cumhal*) ... In the Germanic codes it is mainly in precious metal ... In the Russian codes it is silver and furs.” (op. cit., p. 20)

Payments were made “directly to the victims or their families [and] not to public institutions.” (Hudson 2004, p. 99) The compensations required were quite specific, with different compensations for different offenses. These payments “were established in public assemblies, and the common standards were based on objects of some value which a householder might be expected to possess or which he could obtain from his kinsfolk.” (ibid.) As tribal society did not have writing, a designated *rememberer* was responsible for passing the list of fines to the next generation.

Even though payment of compensation required social consensus on the form of payment, there was no need to settle on a “universal equivalent,” for each specific injury inflicted put a specific debt on the individual (or the family of the transgressor), and each fine was levied in terms of a specific good that was both useful to the victim’s family and more-or-less easily obtainable by the transgressor. Thus, while *wergild* may have been the original source of the practice of measuring indebtedness, it probably could not have directly generated monetary payments because there was little incentive for standardization of the terms.

Nor was there a need for small denominations for payment: If our monetary standards came from the practice of measuring wealth, paying compensations for injuries, or paying bride price, then it is not surprising that the early money units would be large (for example, representing the value of six sheep, or two-months’ worth of grain). This resolves a conundrum that has bothered historians: even the smaller denomination of ancient coins was too valuable to be used for retail trade—indicating they were likely used to pay salaries to “soldiers and sailors,” for wholesale trade, and for paying large debts (such as fines and taxes). This lends further credence to the view that money originates in debts, not as a medium of exchange to resolve the inconveniences of barter.

Hence, Grierson (1975, 1977) rules out the belief that the money of account developed within a market context. Turning to linguistic evidence, Grierson (1977) observes that the Greek word for “price”—*timē*—meant “compensation,” “satisfaction,” and “penalty.” (op. cit., p. 27) He argues that it was a huge intellectual leap to transform the *wergild* fines that were *in kind* to fines measured in a single monetary unit. This would not have been accomplished by agreement among individuals (as in the Robinson Crusoe-Friday story), but rather in a public assembly. One possibility is that with the rise of authorities, they would insist that they would share part of the fine with the victim. As a class society replaced the egalitarian tribal society, the authorities would gain more power to capture the payments. *Wergild* fines may have been transformed into tribute or taxes.

Perhaps the creation of a generalized money of account was not to reduce market transaction costs but was used by rulers to impose obligations on subjects. The practice of paying to *pacify* or eliminate one’s debt for injuries to another may have accustomed the population to the notion of measuring value, and the palace would have had a great incentive to standardize the measure of value (even though neither individuals nor social consensus would have had such an incentive). While the palace could have obtained whatever it needed by imposing *in-kind* taxes with a list of every item it wished and imposing specific taxes on specific producers (for example, a sheep tax on the sheep producer, and so on), imposing a “five *mina*” head tax on each, then using *mina*-denominated palace money to purchase needed items while accepting the same *mina*-denominated palace money in payment of taxes would be a far simpler process.

John Henry locates the origins of money in the rise of inequality within class society. The critical point Henry makes is that the rise of class society and inequality took place alongside the emergence of money, whereby money played a key role in establishing, maintaining, and exacerbating inequality and class division in societies (see Henry 2004). To put it simply, as soon as one witnesses the emergence of money, one observes the rise of class society and socioeconomic inequalities. Money, class society, and inequality came into being simultaneously, so it seems, mutually reinforcing the development of one another. Along with the creation of classes, we find the rise of authorities to protect an unequal distribution of wealth. Originally, these were religious authorities—as in Mesopotamia and Egypt—but religion began to play a lesser role in Greek and Roman civilizations. We turn next to the connections among civilization, the state, and money.

4. Origins of the Money of the State

We have examined three hypotheses about money's origins. While they differ in some details, they have much in common. In this subsection, we will integrate ideas from the three and focus on the link between money and the rise of the state. Much of this is, admittedly, speculative.

However, as discussed, we do have evidence of financial transactions from Mesopotamia from 2500–1200 BC. From this evidence, Michael Hudson concludes that “[t]he kind of general-purpose money our civilisation has come to use commercially was developed by the temples and palaces of Sumer (southern Mesopotamia) in the third millennium BC.” (Hudson 2004, p. 99) With the development of large palace communities, heavy taxes in the form of barley were imposed on producers (initially on villages rather than on individuals). As discussed, Mesopotamia had a dual standard, barley and silver. Although the silver was not coined, the *ruler* (responsible for establishing the measuring units) announced the conversion rate of silver to barley and accepted either in payment of taxes. However, normally producers did not have access to silver, so typically only merchants paid taxes in the form of silver.

The temples played a further role by acting as neutral witnesses, recorders, and enforcers of private transactions (including compensation for damages and payment of bride wealth), and by acting as depositories for grain. At first, these actions would have been recorded on the clay tablets in the grain or cow that they directly represented, with a barley grain fee imposed for the functions provided by the temple. Over time, however, the units became standardized (typically in the name of a grain measuring unit—the *gur*, for example), so that transactions in cows would have been recorded in the money unit, with fees recorded in the same unit (but payable in their equivalents).

As all three approaches argue, rather than originating as a medium of exchange, money developed as an acknowledgment of one's debt—or as a credit, if viewed from the perspective of the creditor. Innes notes that the verb *to pay* has the root meaning *to appease*, *pacify*, or *satisfy* (Innes 1913, p. 392), and he maintains that the *really important* characteristic of money is “the right that it confers on the holder to liberate himself from debt by its means.” (op. cit., p. 393) The *primitive law of commerce*, according to Innes, is “the creation of credits and debts, and their extinction by being canceled against one another.” (op. cit., p. 393) “The presence of a law of debt ... the sanctity of an obligation,” argues Innes, is the foundation of all societies of which we have historical records (op. cit., 391)—that is to say, civilized societies that have developed writing and records of debts.

The rise of state power is associated with the later creation of coins in Lydia (Greece)—which only dates from the 7th century BC. Rather than coming out of market exchange, coinage was a product of a complex political evolution,

and the development of coinage was closely linked to state formation. The rise of state power with a “thoroughgoing monetization of economic activity” required a more “aggressive insistence of the state,” such as “insistence on payments of monetary taxes.” (von Reden 2010, p. 66; see also Kraay 1964) To achieve that, the state would have to “propagate and implement its own commitment to money” (von Reden, p. 66) by using coinage as a means of payment in its own transactions (see also Peacock 2013).

According to Seaford (2004; see also Seaford 2008), the centralized use of coinage for both state expenditures and contributions to the state facilitated the democratic control (by the *polis* in the case of Greece) of the financial system.

... this centralized use of coinage both for expenditure (notably uniform mass payments) and for contributions (liturgies, taxes) must have facilitated – after the fall of the tyranny – the control of the financial system by the democracy, which was thereby able to distance such crucial institutions as the military and the law courts from the influence of personal patronage. (Seaford 2004, p. 99)

As discussed, the money of account long predates the use of precious metal coins. Indeed, evidence suggests that Pheidon’s coins (7th century BC) replaced earlier iron spits (*oboloi*) that had been used as a kind of currency. These had been issued in the grain weight units of account with a stamp to indicate the issuing temple. Coinage was a later development, often with a stamp to indicate the issuer but only rarely (until recently) with a stamp to indicate nominal value. Moving to precious metals seems to have been done to reduce counterfeiting, since scarce metals would be harder to obtain. (Heinsohn and Steiger 1983) The precious metal *veil* that has clouded monetary thought ever since apparently resulted from this purely technical consideration. As Innes argues,

What has really happened is that the government has put upon the pieces of gold a stamp which conveys the promise that they will be received by the government in payment of taxes or other debts due to it ... In virtue of the stamp it bears, the gold has changed its character from that of a mere commodity to that of a token of indebtedness. (Innes 1913, p. 402)

Similarly, Theodor Mommsen argued that in the case of Roman coinage,

It may be regarded as a law which gives the ... piece of metal its conventional value by legal decree, quite irrespective of whether the effective value corresponds with this or not. In this, so to speak, statutory validation, the coin of the realm ... is already enshrined in republican law: only this coinage is money - all others are commodities of trade. (quoted in Heinsohn and Steiger 1983, p. 22)

According to Heinsohn and Steiger:

In the ancient world, at all events, there was a full awareness of this gold fog obscuring the true nature of money. Aristotle, for instance ... writes: "In some respects, however, money is a pure sham, a creature of convention established in law". (ibid., p. 23)

To sum up, early money units appear to have been derived from weight units that may have developed from the practice of *wergild*. Palaces created money units to standardize the payment of debts, and later, taxes. The use of money in private transactions derived from debts encouraged by the palaces, which could record and enforce private transactions. Once a money tax or tribute was levied on a village, and later on individuals, the palace would be able to obtain goods and services by issuing its own money-denominated debt, initially in the form of clay tablets and later wooden tallies (see discussion below). Coins came much later, but were, like the tallies, evidence of the Crown's debt. The use of precious metals in coins was adopted—largely—to reduce counterfeiting.

5. Conclusion on the Alternative Approaches to Money's Origins

As mentioned at the outset, these three alternatives are not mutually exclusive, and each probably captures important elements of the origins of money. Nor do we need to choose a single origin story—money may well have multiple histories—across space and time—and these three share important elements. All reject the barter story and point to a nonmarket source for money's origins. All emphasize the money of account function—the way that credits and debts are recorded. Each also recognizes that an economy that uses money is different, and although money dates back thousands of years, it represents a deviation from the way that humans lived for many more thousands of years before money appeared on the scene. Authorities play a role, and so does private property. In these approaches, money is also linked to interest in a manner that is more consistent with Keynes (where interest is a monetary phenomenon) than with neoclassical theory (where interest is linked to physical productivity). These are themes to which we will return.

IMPLICATIONS FOR OUR UNDERSTANDING OF THE ROLE OF THE STATE AND THE NATURE OF MONEY

As discussed, the orthodox story of money's origins focuses on trade in markets, with money serving as a medium of exchange. An evolutionary process selected precious metal to serve as money, and eventually, the metal would be coined to ensure consistent value. Only later would fiat money and credit play an important role in markets—and regrettably so, as money's value is eroded because it is not backed by precious metal. In this section, we examine coins

in detail, arguing they generally play a subsidiary role in markets, although they have been important for government finance—albeit not as orthodoxy has supposed.

Historical evidence suggests that most commerce from the very earliest times was conducted on the basis of credits and debits—not through immediate payment using a medium of exchange. Innes writes of the early European experience: “For many centuries, how many we do not know, the principal instrument of commerce was neither the coin nor the private token, but the tally.”¹ (ibid., p. 394). This was a “stick of squared hazel-wood, notched in a certain manner to indicate the amount of the purchase or debt”, created when the buyer became a debtor by accepting a good or service from the seller who automatically became the creditor. (ibid.) “The name of the debtor and the date of the transaction were written on two opposite sides of the stick, which was then split down the middle in such a way that the notches were cut in half, and the name and date appeared on both pieces of the tally.” (ibid.) The split was stopped about an inch from the base of the stick so that one piece, the *stock*, was longer than the other, called the *stub* (also called the *foil*). The creditor would retain the stock (from which our terms capital and corporate stock derive) while the debtor would take the stub (a term still used, as in *ticket stub*) to ensure that the stock was not tampered with. When debtors retired their debts, the two pieces of the tally would be matched to verify the amount of the debt.

Wooden tallies were not the only records, as there was nothing unique about hazelwood (indeed, it appears to have been used because it was common in England and Northern Europe). As discussed above, the earliest records of debts come from Babylonia, on clay *shubati* (*received*) tablets; these indicated a quantity of grain, the word *shubati*, the name of the person from whom received, the name of the person by whom received, the date, and the seal of the receiver or of the king’s scribe (when the king was the receiver). Unlike the wooden tally, these tablets would not be split to give the debtor a stub. The verification problem was solved in two ways: the tablets were either stored in temples where they would be safe from tampering, or they were sealed in cases that would have to be broken to reach them. All the inscriptions listed above would be repeated on the case, but the enclosed tablet would not contain the name and seal of the receiver. Thus, if the case were broken, the tablet would not be complete. Only when the debt was repaid would the case be broken (allowing the debtor to observe that the inscription on the case matched that of the enclosed tablet).

As discussed in Chapter 3, tallies eventually could circulate as “transferable, negotiable instruments.” One could deliver the stock of a tally to purchase goods and services, or to retire one’s own debt. “By their means all purchases of goods, all loans of money were made, and all debts cleared.” (Innes 1913,

p. 396). A merchant holding tally stocks of customers could meet with a merchant holding tally stocks against the first merchant, clearing tally stub debts by delivery of the customers' stocks. In this way, according to Innes, great *fairs* were developed to act as *clearing houses* allowing merchants "to settle their mutual debts and credits"; the "greatest of these fairs in England was that of St. Giles in Winchester, while the most famous probably in all Europe were those of Champagne and Brie in France, to which came merchants and bankers from all countries." (ibid.) Debts were cleared "without the use of a single coin"; it became common practice to "make debts payable at one or other of the fairs," and "[a]t some fairs no other business was done except the settlement of debts and credits," although retail trade was often conducted at the fairs. While conventional analysis views the primary purpose of the medieval fairs as retail trade, Innes postulates that the retail trade originated as a sideline to the clearing house trade.² He also notes that clearing house fairs were also held in ancient Greece and Rome, and in Mexico at the time of the conquest.

Even if one accepts that much or even most trade took place based on credits and debts, this does not necessarily disprove the story of the textbooks. Perhaps coins existed before these tallies (and other records of debts), and surely the coins were made of precious metals? Perhaps the debts were made convertible to coin; indeed, perhaps such debt contracts were enforceable only in legal tender coin? If this were the case, then the credits and debts merely substituted for coins, and net debts would be settled with coin—which would not be inconsistent with the conventional story. There are several problems with such an interpretation, however.

First, the recorded debts (in the form of clay tablets) are at least 2000 years older than the oldest known coins. It seems unlikely that clay tablets would outlast coined precious metal. Second, as mentioned earlier, it has long befuddled economic historians that the denominations of the early precious metal coins were generally too high to have been used in everyday commerce. For example, the earliest coins were electrum (an alloy of silver and gold) and the most common denomination would have had a purchasing power of about ten sheep, so that "it cannot have been a useful coin for small transactions." (Cook 1958, p. 260) They might have sufficed for the wholesale trade of large merchants, but they could not have been used in day-to-day retail trade. Furthermore, the reported nominal value of coins does not generally appear to be closely regulated by precious metal content. It is also unlikely that coins would have been invented to facilitate trade, for "Phoenicians and other peoples of the East who had commercial interests managed satisfactorily without coined money" for many centuries. (ibid., p. 260) Indeed, the introduction of coins would have been a less efficient alternative in most cases.

It is also useful to repeat that coins did not have a denomination stamped on them until relatively recently. Their nominal value would be set by

announcement—for example, the medieval *town-crier* would announce the value—and would be changed from time to time. This would have been a futile activity if the *real* value were determined by precious metal. But in practice, the kings (or other authorities who issued them) would set the value at which the coins would be accepted in payment. By *crying up* or *crying down* coins, the crown would increase or reduce the coins' value in payment of taxes and other obligations.

Finally, while we are accustomed to a small number of types of coins (always issued by the government, with perhaps one coin for each denomination), the typical case until recently was a plethora of coins, sometimes including many with the same face value but different exchange values, issued by a wide variety of merchants, kings, feudal lords, barons, ecclesiastics, and others. Indeed, “in [feudal] France there were besides the royal monies, eighty different coinages ... each entirely independent of the other and differing as to weights, denominations, alloys, and types [and] twenty different monetary systems.” (Innes 1913, p. 385) According to MacDonald, in Merovingian Gaul there were “1200 different moneyers,” the great majority of whom were private individuals; this “epoch of private coinage” seems to have been “brought to an end by Pepin and Charlemagne.” (MacDonald 1916, pp. 29–35)

Note that the textbook story relies on the choice of a particular precious metal by “common consent” to be used as money to reduce the transaction costs of barter. However, in reality, the poor consumer (if such existed) was faced with a tremendous number of coins of varying weight, denomination, alloy, and fineness with which they would not be able to cope. Indeed, it is difficult to believe that the typical members of these societies would be more able to assess the value of a coin than they would be able to assess the value of, say, a cow. Rather than reducing transaction costs by using precious metals, it would probably have reduced transaction costs to use cows! And it does no good to argue that cows are less divisible, for as noted above, the precious metal coins were too valuable to have been used in daily transactions anyway. (Although, as we know, coins could be cut into “bits” from which our term “two bits,” a quarter of a dollar—the cost of a haircut at the turn of the 20th century—derives.)

That many were not used in frequent transactions is evidenced by “the excellent state of preservation in which they are usually found.” (Grierson 1965, p. 536) Estimated “wear and tear” on coins (especially gold coins) in circulation is quite high—perhaps 1 percent per year (Munro 1979, pp. 181–2)—but “Carolingian coins seem to have circulated surprisingly little.” (Grierson 1965, p. 536) Finally, Grierson notes that it was frequently necessary to impose “legislation forcing people to use coin; if they refused it they laid themselves open to severe penalties, a heavy fine if they were free men or a flogging if they were unfree.” (ibid.) This hardly seems consistent with the textbook story

of “common consent”—and would be irrational if coins contained precious metal worth the nominal value of the coin!

It is also difficult to understand why precious metal coins were normally worth more than would be dictated by their precious metal content if it is true that the value of the precious metal determines the value of the coin. Clearly, if the nominal value of the coin were below the relative value of the precious metal contained therein, the coin would be removed from circulation to be used as metal. But, given the costs of coinage, if the mint were to issue coins whose value was little more than that of the embodied metal, this would provide very little purchasing power to the mint. While the textbook story argues that paper credit developed to economize on precious metals, we know that metal coins were a relatively late development. In other words, lower-cost alternatives to full-bodied coin were already in use. Surely hazelwood tallies or clay tablets had lower non-monetary value than did precious metals. Thus, it is unlikely that metal coins would be issued to circulate competitively (for example, with hazelwood tallies) unless their nominal value was above the value of the embodied precious metal.³

What then are coins, what are their origins, and why are they accepted? Coins appear to have originated as *pay tokens* (in Knapp’s colorful phrase), as evidence of debt. It is possible that they were derived from medals that were common in some traditional societies. The earliest *coins* then, may have been nothing more than gifts with an imprint to signify the giver; it is conceivable that these were given to recognize a personal debt to the receiver. As noted earlier, the first coins were struck by authorities, probably by Pheidon of Argos about 630 BC. (Cook 1958, p. 257) Given the large denomination of the early coins and uniform weight (although not uniform purity—which probably could not have been tested at the time), Cook argues that “coinage was invented to make a large number of uniform payments of considerable value in a portable and durable form, and that the person or authority making the payment was the king of Lydia.” (ibid., p. 261) Further, he suggests, “the purpose of coinage was the payment of mercenaries.” (ibid.)⁴

This thesis was modified “by Kraay (1964) who suggested that governments minted coins to pay mercenaries only to create a medium for the payment of taxes.”⁵ (Redish 1987, pp. 376–7) Crawford has argued that the evidence indicates that the use of these early coins as a medium of exchange was an “accidental consequence of the coinage,” and not the reason for it (Crawford 1970, p. 46). Instead, Crawford argued that “the fiscal needs of the state determined the quantity of mint output and coin in circulation”; in other words, perhaps coins were intentionally minted from the beginning to provide *state finance*. (ibid.) So, early governments understood that “[m]inting and taxing were two sides of the same coin of royal prerogative.” (Davies 1997, p. 146)

Similarly, Innes argued that “[t]he coins which [kings] issued were tokens of indebtedness with which they made small payments, such as the daily wages of their soldiers and sailors.” (Innes, 1913, p. 399) This explains the relatively large value of the coins, which were not meant to provide a medium of exchange but rather were used to settle the state’s debt to “soldiers and sailors.” The coins were then nothing more than “tallies” as described above—evidence of government debt—and not deserving of the inordinate concern shown by modern economists. And relative to the quantity of hazelwood tallies and other forms of money, the quantity of coins was quite small:

[i]ndeed so small was the quantity of coins, that they did not even suffice for the needs of the Royal household and estates which regularly used tokens of various kinds for the purpose of making small payments. So unimportant indeed was the coinage that sometimes Kings did not hesitate to call it all in for re-minting and re-issue and still commerce went on the same.⁶ (Innes 1913, p. 389)

Let us step back for a moment and ponder the implications. In our view, coins are mere tokens of the Crown’s debt, a small proportion of the total *tally*.

Just like any private individual, the government pays by giving acknowledgments of indebtedness – drafts on the Royal Treasury, or some other branch of government. This is observed in medieval England, where the regular method used by the government for paying a creditor was by “raising a tally” on the Customs or some other revenue-getting department, that is to say by giving to the creditor as an acknowledgment of indebtedness a wooden tally. (Innes 1913, p. 397–8)

The *tallia divenda* developed to allow the king to issue an exchequer tally for payment for goods and services delivered to the court.⁷ But why would the Crown’s subjects accept hazelwood tallies or, later, paper notes or token coins?

The government by law obliges certain selected persons to become its debtors. It declares that so-and-so, who imports goods from abroad, shall owe the government so much on all that he imports, or that so-and-so, who owns land, shall owe to the government so much per acre. This procedure is called levying a tax, and the persons thus forced into the position of debtors to the government must in theory seek out the holders of the tallies or other instrument acknowledging a debt due by the government, and acquire from them the tallies by selling to them some commodity or in doing them some service, in exchange for which they may be induced to part with their tallies. When these are returned to the government Treasury, the taxes are paid. (ibid., p. 398)

Innes went on to note that most revenues collected by inland tax collectors in England were in the form of the exchequer tallies:

[p]ractically the entire business of the English Exchequer consisted in the issuing and receiving of tallies, in comparing the tallies and the counter-tallies, the stock and the stub, as the two parts of the tally were popularly called, in keeping the accounts of the government debtors and creditors, and in cancelling the tallies when returned to the Exchequer. It was, in fact, the great clearing house for government credits and debts. (ibid.)

Each taxpayer did not have to seek out individually a Crown tally, for matching the Crown's creditors and debtors was accomplished "through the bankers, who from the earliest days of history were always the financial agents of government." (Innes 1913, p. 399) That is, the bank would intermediate between the person holding Crown debt and the taxpayer who required Crown debt to pay taxes.

The exchequer began to assign debts owed to the king whereby "the tally stock held in the Exchequer could be used by the king to pay someone else, by transferring to this third person the tally stock. Thus, the king's creditor could then collect payment from the king's original debtor." (Davies 1997, p. 150) Further, a brisk business developed to *discount* such tallies so that the king's creditor did not need to wait for payment by the debtor. (See the discussion in Chapter 3.) Note, also, that the use of the hazelwood tallies continued in England until 1826. Ironically, the tallies went out in a *blaze of glory*, or of ignominy, depending on one's point of view. After 1826, when tallies were returned to the exchequer, they were stored in the Star Chamber and other parts of the House of Commons. "In 1834, in order to save space and economize on fuel it was decided that they should be thrown into the heating stoves of the House of Commons. So excessive was the zeal of the stokers that the historic parliament buildings were set on fire and razed to the ground." (Davies 1997, p. 663)

The inordinate focus of economists on coins (and specifically on government-issued coins), market exchange, and precious metals, then, appears to be misplaced. The key concept is debt, and specifically, the ability of the state to impose obligations—today, mostly tax debts—on its subjects or citizens. Once it has done this, it can choose the form in which the tax can be paid. While the government could, in theory, require payment in the form of all the goods and services it requires, this would be quite cumbersome. Thus, it becomes instead a debtor to obtain what it requires (and note that this is no different from the way in which most buyers become debtors), and issues a token (hazelwood tally, coin, paper note, or, today, entries onto balance sheets of banks) to indicate the amount of its indebtedness; it then accepts its own debt in payment to retire tax liabilities. Certainly, its token of debt can also be used as a medium of exchange (and means of debt settlement among private individuals), but this derives from the government's ability to impose obligations and its willingness

to accept its own debts. Indeed, if one has a tax liability but is not a creditor of the Crown, one must offer things for sale to obtain the Crown's debts.

In the next chapter, we will explore later developments of credit and the rise of banking. We will return to the State Money approach in Chapters 4 and 8.

NOTES

1. Davies (1997) also notes the *ancient* origins of tallies and quotes Anthony Steel to the effect that "English medieval finance was built upon the tally." (Davies 1997, p. 147) The word tally seems to have come from the Latin talea, which means a stick or a slip of wood; notches in sticks had long been used for recording messages of various kinds. (Davies 1997, p. 147) Note that one of the most common "notches" was the score, which indicated 20 pounds; a one-pound notch was a small groove the size of a barley grain—see the discussion below.
2. Some merchants may have brought goods to the market to use to settle accounts, with retail trade developing from this practice. Admittedly, the view expounded by Innes is controversial and perhaps too extreme. What is important and surely correct, however, is his recognition of the importance of the clearing house trade to these fairs.
3. It is often asserted that coins were invented to facilitate long-distance trade (as precious metal coins would have high value relative to weight). As Grierson notes, "The evidence, however, is against the earliest coins having been used to facilitate trade of such a kind, for the contents of hoards point overwhelmingly to their local circulation." (Grierson 1977, p. 10)
4. Grierson also advances this thesis: "The alternative view is that since coins were issued by governments – the supposed issue of the earliest coins by merchants is unproven and unlikely – it was administrative rather than economic needs they were intended to serve. Such needs would have included the payment of mercenaries ..." (Grierson 1977, p. 10)
5. Crawford suggests that "[c]oinage was probably invented in order that a large number of state payments might be made in a convenient form and there is no reason to suppose that it was ever issued by Rome for any other purpose than to enable the state to make payments ..." (Crawford 1970, p. 46) Further, "[o]nce issued, coinage was demanded back by the state in payment of taxes." (ibid.)
6. Recoinage would be a strange activity if the value of the coin were determined by the value of the embodied precious metal. The modern equivalent is to call in the coins and knock three zeros off the reissue—an activity that is easy to explain in the case of a fiat money.
7. The wooden tallies were supplemented after the late 1670s by paper *orders of the exchequer*, which in turn were accepted in payment of taxes. (Grierson 1975, p. 34)

3. Credit money and the rise of banking, central banking, and public money

As we saw in Chapter 2, the evidence suggests that the money of account was created by authorities for internal record keeping and then adopted as the *nominal* measuring unit for debt and credit. Records of indebtedness could be used in trade and the payment of one's own debt. Strong central authorities could spend using their own records of indebtedness, redeemed in payment of fees, fines, taxes, and tithes—as Innes argues. This chapter will focus on early credit money, the rise of banking, the creation of central banks, and the role they played in developing a monetary system based on public money. While both public money and private money have existed for 4000 years (*at least*, as Keynes claimed), the money of account in which they are denominated has almost always been chosen by authorities. Still, the circuits of circulation of each have at times been quite separate, with the relative importance of them contested and shifting. With the rise of capitalism and the modern monetary system, the circuits became closely integrated, with the state's money at the top of the hierarchy.

We begin with the credit arrangements of early Islamic society and then move on to early Western Europe. While it is commonly assumed that commodity money preceded the development of credit money, the documentary evidence tells a different story: that credit relations originated earlier. Moreover, the evidence contradicts the common view that precious metal determined the value of coined money—what has been called *theoretical commodity money* by Ingham. This probably has never been the typical arrangement. Instead, there have been instances of what we can call—following Ingham—*practical commodity money*, in which the issuer of coins sets their value relative to the money of account and maintains a relatively constant quantity of embodied precious metal in the coins.

This does not mean the coin's value is determined by the value of the metal, but rather that the monetary unit determines the value of that quantity of metal. As Keynes put it in 1913, a *rupee* is a promissory note printed on silver; Ingham insists, “*Money*ness is conferred on a substance or form by the unit of account.” (2004a, p. 212) This is not to deny the importance of precious metal and “bullionism” in the development of the modern monetary system,

however. As we will see, maintaining “full bodied” coins in England increased the trust that eventually led to the integration of private and public money into a unified monetary system effectively under the auspices of the central bank.

The modern monetary system was created with great difficulty—in large part because, after the fall of Rome, governments were typically weak. They frequently saw private credit instruments and banking as competitors and threatening. The mistrust was mutual as lending to the crown could be risky and even dangerous. By happenstance, it was the default by Charles II in 1672 that set off a series of reforms that finally produced the modern monetary system that allowed capitalism to develop. Strangely enough, a weakened sovereign combined with the maintenance of a strong metallic coin promoted the development of private banking governed by a national central bank while increasing the fiscal power of the state. State money dominance and currency sovereignty emerged from this arrangement—with private credit money fully integrated into the national monetary system and obscuring that dominance.

DEFINITIONS

We first define several terms that will be used throughout this chapter. Most of these signify various financial instruments or practices that may not be familiar to modern readers.

A bank note is paper currency issued for circulation by banks and may be convertible on demand into other forms of currency such as state-issued money (coins or paper money), or into privately issued liabilities (for example, Bank of England notes). However, convertibility is not a necessary characteristic of bank notes—they have often circulated even when not convertible.

Redemption will be defined as the return of issued liabilities to the issuer. There are two kinds of redemption: redemption for conversion and redemption for payment. For example, when bank notes return to the issuing bank, redemption has occurred. Returned notes may be deposited (changing the bank’s liability from a note to a deposit), may be used to repay debt to the bank (reducing the bank’s asset and liability), or (in the case of convertible notes) may be redeemed for currency issued by the government. The last operation causes a loss of reserves so that this form of redemption is undesirable from the point of view of banks. It is also unnecessary—banknotes may circulate for long periods before redemption occurs.

Redemption of a bank’s deposits occurs as depositors draw upon their accounts: a check received in payment will be presented to another bank and reserves will be lost through redemption during the check-clearing process. The bank may attempt to recapture these reserves by attracting deposits (or through the issuance of some other form of liability).

The bill of exchange is used to purchase goods on credit that will be retired later when the goods are sold. Typically, the issuer of the bill (borrower) would receive the goods and promise to redeem the bill at a specified location, in a specified currency, and on a specified date. Interest could be included in the bill of exchange, as the borrower would promise a greater quantity of currency on the deferred date. The acceptor of the bill (lender) would find a way to return the currency received on the deferred date to the acceptor's domicile. Rather than transporting currency (which was often foreign), the acceptor would prefer to have an agent purchase goods or another bill of exchange—this time drawn on the domicile of the original lender. Originally, bills of exchange were used for long-distance trade but they evolved into a credit instrument used in diverse activities.

Bills of exchange could be used to evade usury laws, as interest could be hidden in exchange rates. Often, a bill of exchange would be issued in one currency, with payment later in another currency. By using a disadvantageous exchange rate, the issuer was effectively paying interest. Alternatively, the bill would include a penalty for late payment, with the acknowledgment that payment would be made after the scheduled date—equivalent to charging interest.

Eventually, bills of exchange became negotiable: the original lender might sell the bill of exchange rather than wait to receive payment at the specified location. As a secondary market for bills developed, they could be used as a medium of exchange and to pay debts. Bills of exchange could be *endorsed* by respected individuals to increase their acceptability; bills with a large number of impressive endorsements could be said to be gilt-edged—"as good as gold".

A *giro* is defined as a *payment society* in which the members agree to accept credit instruments of one or more of the members as media of exchange and as means of payment. The wool trade of England in the 14th century serves as an example of a *giro*. English wool exporters (*staplers*) required immediate payment in cash of one-third and accepted bills of exchange issued by Flemish importers (*drapers*) for the remaining two-thirds. The *drapers* would sell the wool and buy a bill of exchange issued by a Flemish merchant payable in London. The merchants would export goods to London to be sold to retire the bill of exchange. The *staplers* would finally be paid when these imported goods were sold at the London fairs. Since the *staplers* had usually bought the wool from producers by issuing bills of exchange, they would now be able to retire their own bills. Thus, the bill of exchange allowed the wool to move from the producer, to the *stapler*, and to the *draper*. This circuit gave rise to the bill issued by the Flemish merchant and to another circuit of goods going in the reverse direction: from Flanders to London. After the development of the negotiable bill of exchange, the size of the *giro* and the complexity of the circuits would be increased.

The functions of a bank can be characterized as the *Lombard* business and the *giro* business. (Knapp 1924, p. 129) In the *Lombard* business, banks provide money now for the promise of money later. This may take the form either of a discount of eligible paper (for example, a bill of exchange) or of an advance on certain kinds of goods that will be sold. In the first case, the bank earns interest by buying the paper at a price below face value, and in the second, it explicitly charges a rate of interest on the loan.

Banks also facilitate circulation by becoming part of the *giro*. The accounts of the bank are used to transfer money (as a unit of account) between debtors and creditors. The bank may also act as a guarantor, endorsing bills of exchange to circulate within the *giro*. The business of a bank can be expanded if it can get its customers to accept bank debt in payment. The bank can increase its profits by issuing notes in its *Lombard* business so that the volume of its discount business is not tied to the quantity of minted coin (or other forms of money) held in its vaults. (Knapp 1924, p. 131) From the earliest period, banks issued liabilities, such as notes, to make loans. As the notes become generally accepted, the size of the *giro* is increased and the ability of the bank to expand is enhanced.

Eventually, the bank debt can displace the debt of the other members of the *giro*, so that the bank becomes the apex through which most other debts must pass. That is, the bank issues its notes in loans and accepts its notes in deposits. Bank customers use bank notes between themselves because the notes can be used to settle debts with the bank. Neither convertibility into other forms of money (for example, state currency), nor circulation outside the *giro* is a requirement for circulation within the *giro*. (Knapp 1924, p. 135) However, the possibility of use outside the *giro* will clearly become important as the scale of potential payments outside the *giro* grows, and convertibility may be required before the *giro* can expand. The state can become a member of the bank *giro*, encouraging the process through which bank liabilities supplant other private liabilities in the *giro*.

Commodity money is usually defined as currency whose value is determined by an objective measure—usually the weight of precious metal. As discussed in Chapter 2, the weight unit used to measure grains of barley or wheat was borrowed as the first money unit, and equivalent weights of gold, silver, bronze, and iron were later used. Thus, minted coins containing a specific quantity of a precious metal are commodity money.

Fiat money is currency whose value is purely nominal. The term nominalism specifies that a coin's value is set by proclamation or law, not by the quantity of metal contained in the coin—even when precious metals may be used in its production. As we will argue in this chapter, through most of history, the quantity of precious metal in coins varied greatly—even those with the same value in terms of the money of account—and the nominal value was

normally significantly above the value of the embodied precious metal (which often would have been difficult to determine). On the other hand, some issuers did maintain a relatively constant amount of precious metals in their coins, including Britain for several centuries. However, we will argue that even in that case, the value of gold did not determine the money value of the coins—rather, it was the other way around: the issuer set a money value on the metal used in coinage.

Legal tender will be defined as any form of money which the state has officially recognized for payment of debts. In most cases, any form of money accepted in payment of obligations to the state (today, mostly taxes, although fees, fines, tithes, and tribute have all been used at various times) may be called legal tender. By accepting private liabilities (such as bank notes or checks) in payment of debts to the state, the state becomes a member of the private *giro*.

CREDIT MONEY WITHOUT BANKS: THE CASE OF ISLAMIC SOCIETY

The medieval Islamic world serves as a useful example for an exposition of the use of sophisticated credit arrangements in the absence of banks, deposits, and explicit interest. Islamic law provided severe sanctions against interest-bearing loans, speculative transactions, certain forms of delayed payment, and other commercial and exchange operations. (Udovitch 1979, pp. 256–7) The prohibition against usury and interest “precluded the coalescence of two financial practices ... the generation of capital by means of numerous deposits on the one hand, and moneylending and the provision of credit on the other.” (Udovitch 1979, p. 259) Under Islamic law, a depositor can neither earn interest nor be charged a fee. Furthermore, anyone who accepts a deposit must hold it for safekeeping and return it on demand. Thus, it was impossible for so-called intermediaries to develop that would accept deposits and loan them at interest.

However, such proscriptions did not prevent the development of sophisticated market activities normally associated with societies with evolved banking institutions, for “as early as the late 8th century, credit arrangements of various types were an important feature of both trade and industry in the Near East.” (op. cit., p. 260) Merchants bought and sold goods on credit—whether for local or international trade. One of the techniques utilized was the *credit partnership* in which “people form a partnership without any capital in order to buy on credit and then sell.” (op. cit., p. 263) It was called a *partnership of good reputations* because credit was advanced on the basis of the reputations of the partners. The partnership could thus obtain commodities on credit, hoping to sell them later at a higher price so that they could retire their debt and retain profit. The sanctions against interest charges were avoided by quoting two prices: one for immediate payment and one for future payment. Islamic

law permitted a premium for deferred payment (which is implicitly an interest rate) because the merchant faced some risk that the partnership would default.

Bills of exchange, *suftajas*, were widely used as a means of deferred payment. However, unlike the (later) Western bill of exchange, the *suftaja* did not involve the exchange of currencies and did not develop into a negotiable instrument. Paper notes (*ruqcas*) were issued by well-known merchants and “could be used not only in commerce but even to pay government toll collectors.” (Udovitch 1979, p. 270) These were also used much like modern checks to settle petty debts. Most merchants engaged in banking activities such as: “money changing, issuing letters of credit, accepting deposits, [and] acting as a clearinghouse for payments ...” (op. cit., p. 271). Merchants allowed overdrafts by good customers and accepted promissory notes in lieu of cash payments. Thus, they provided checking accounts and made loans by issuing their own liabilities even though they were prohibited from loaning the deposits of depositors.

Note that medieval Islamic society was not a *monetary economy*. Medieval Arabic has no words for banks, banking, or money. Credit was closely tied to commercial activities and was advanced based on social-personal relationships among merchants and their customers. Most production was undertaken for nonmarket use, so “the scale of economic activities was necessarily restricted to numerous small, even intimate, circles. The possibility of expansion into a larger, more cohesive structure was precluded by the comparatively narrow social basis on which economic life was conducted.” (op. cit., p. 273) The development and organization of financial activity remained limited.

Thus, the case of Islamic society supports the view that credit money can be created on demand, as a private liability to give purchasing power to the borrower. The credit advanced by merchants represented liabilities of the merchants and was constrained by their *good reputation*. The orthodox story in which money creation occurs when goldsmiths realize they need only hold a small reserve against notes issued cannot apply to Islamic society, where this intermediary function was prohibited by law.

PRE-MODERN FINANCIAL INSTITUTIONS IN WESTERN EUROPE

For the purposes of this chapter, we define ancient society to cover the period up to the fall of Rome, while the Middle Ages will cover the period from ancient society through the twelfth or thirteenth centuries, and the premodern period will cover the period from the Middle Ages to the modern period, which begins in the late 17th century (when capitalism begins to develop).

As discussed above in Chapter 2, money was first developed as a unit of account probably at least 5000 years ago, perhaps for internal record-keeping within temple communities. The ancient Near East developed a variety of

monetary institutions and financial instruments between 3000 and 500 BC. In the ancient, middle, and premodern periods, there was typically a small monetized sector within a large nonmonetary economy. The barley (or other grain) unit was used as the basis for the first moneys of account, with silver (or other precious metals) later substituted on a weight-equivalent basis. The financial instruments in use were primarily trade and consumer credit. Credit was used to finance inventories (especially in foreign trade), consumption by the aristocracy, or provision of seed corn to producers. The major suppliers of credit included temples, private merchants and landowners, and royal treasuries. They loaned their own reserves of grain, money, or other goods.

Claims against financial institutions were usually of less importance until near the end of the premodern period (after the 13th century in Italy and Spain, and later in England). Interest rates were typically between 20 and 33 percent, although interest rates ran as high as 50 percent on consumer loans. Except in the latter case, the interest rate seems to have been fixed according to the typical rent received from land, and also to ease calculation. In general, the extent of foreign trade serves as a good indication of the degree of development of financial instruments. That is, those societies that were more involved in foreign trade had a higher level of development of financial institutions due to the need to finance goods in the process of transport.

Money in Greece and Rome

We first examine ancient society, taking Greece and Rome as examples, before turning to financial practices during the Middle Ages. According to Seaford, “The earliest evidence for Greek money comes from the time around 600 BC, for instance, in the sums specified as compensation for injury (or fines) in the legislation of Solon.” (Seaford 2008, p. 52) This would appear to be consistent with the belief that money originated from Wergild payments, as Grierson proposed. Still, the scope for the private creation of money was probably limited by the narrow field of private contractual relations. Most production likely remained outside the market. While government spending was significant, it relied on the use of coins—and there is some debate over whether this constrained government spending, although it appears that debasement (reducing the quantity of precious metal in a coin) was used to increase the coinage. When that occurred, the legal records indicate that nominalism prevailed in judgments.

The first coins were produced in Lydia in the 7th century B.C. and were made of a mixture of gold and silver (electrum). However, over the past few decades, scholarship has reduced the former emphasis on coins and provided ample evidence from the literature of the time that credit money played a significant role in the economy—not only for trade in goods but also for purchases

and construction of expensive dwellings. Furthermore, scholars have emphasized the importance of nominalism—while coins contained precious metals, payment of debt using coins took place at the nominal value of the coin at the time of payment.

Unfortunately, the data that exist do not make it possible to be certain of the importance of the use of coins versus credit in transactions. Still, even in early Greece, credit money may have dominated. As Cohen argues:

By the mid-4th century, or perhaps even earlier, Athens was making use of a bronze coinage whose metallic value was far less than its nominal worth, thus constituting a form of ‘unbacked fiat money’. But the dominant form of non-commodity money in the modern world—as at Athens—has been credit money, which in turn consists of bank money (that is, bank deposits more or less broadly defined) and debt-related money in all its manifestations (as a practical matter principally credit generated by sales transactions). (Cohen 2008, p. 68)

Cohen insists that “the money supply in 4th-century Athens was in fact strikingly elastic, since it could be—and was—substantially increased through the provision of credit by merchants and through the banks’ creation of non-coin money via deposit accounts and other mechanisms.” (Cohen 2008 p. 67) Consumer and business loans were financed by banks:

Demosthenes, for example, defines the *trapeza* as “a business operation producing risk-laden revenues from other people’s money”, pithily focusing on the making of loans (“risk laden revenues”: *prosodous epikindynous*) and the receipt of deposits (‘other peoples’ money’) ... (op. cit., p. 68)

The Roman Empire was the largest economic unit in the Western world until the 19th century. Historians had thought that coins dominated a relatively small monetized sector (including the slave trade—with estimates that slaves amounted to between 10 percent and 20 percent of the population of the empire). Thus, Goldsmith argued that “Augustan Rome ... had no fiduciary money, no credit creation, no negotiable instruments, no business loans except bottomry loans ... and no long-term partnerships or corporations” and his focus is largely on the use of precious metal. (Goldsmith 1987, p. 36) More recent scholarship, however, emphasizes the ubiquity of Roman credit and Roman law that favored nominalism. For example, in the case of precious metal coins deposited *in sacculo* (enclosed in a sealed sack) for safekeeping, those same coins had to be returned to the depositor. However, if one was owed a sum of money, one would have to accept payment in any officially sanctioned coins worth that sum. In other words, nominalism was the law of the land.¹

Harris (2008) notes that the argument that Rome did not have credit is highly implausible given the historical accounts of transactions involving

many millions of *sesterces* that would have required tons of coins to make the payment—and historians have not been able to find records of any such payments made in bullion. Instead, he finds, “There is plenty of evidence that rich and poor alike casually and normally bought everyday items on credit, in many different environments.” (Harris 2008, p 178) The range of recorded payments of debts includes:

maritime loans, *cheirographa* (unsecured loans, presumably), *parathekai* (loans secured by portable objects), first and second mortgages, and most interestingly of all debts concerning sales (*kat’onas*) that were in the form of a *homologia* (a legal document acknowledging the receipt of a loan) ... (op. cit. p. 183)

... debt was in fact the lifeblood of the Roman economy, at all levels. The normality of *nomina* (i.e. outstanding loans) among the assets of the rentier class ... were a completely standard part of the lives of people of property, as well as being an everyday fact of life for great numbers of others... Bankers were also able to make payments at a distance, in other words without the direct use of coins, which meant that other people (bankers too, presumably) accorded them credit. (op. cit., p. 187)

To conclude, even in the ancient world, credit money played an important role, even though it did not lead to a *monetary* system in which production for sale on the market for money is the norm. When Rome fell in 476 AD, the size of the monetized sector declined because the feudal mode of production, which was not based on private property, had less need for money for internal relations. However, this does not mean that money and financial relations disappeared, for they were still necessary for external relations. Indeed, financial practices continued to evolve over the next thousand years—largely for long-distance trade.

Money in Medieval Europe

With the fall of Rome, credit needs declined, but old Roman coins continued to be used for trade. For several centuries, there was no central monetary authority. Feudal lords often issued their own currency, including leather “coins.” Other issuers included merchants, barons, kings, and ecclesiastics. For example, in feudal France “there were beside the royal monies, eighty different coinages ... each entirely independent of the other and differing as to weights, denomination, alloy and types [and] twenty different monetary systems.” (Innes 1913, p. 385) In Merovingian Gaul, there were “1200 different moneyers,” the great majority of whom were private individuals, with this “epoch of private coinage brought to an end by Pepin and Charlemagne.” (MacDonald 1916, pp. 29–35)

Charlemagne (768–814) “decreed a common money of account, derived from the Roman system” based on 240 pence to the pound, divided into 20 shillings—a system that persisted until the late 18th century across Europe. (Ingham 2004a, p. 189) But the money of account he created did not correspond to any of the actual minted coins in use. This has led to confusion among historians about the use of an “imaginary money” for accounting purposes, but with no “real money” coins based on that accounting. In truth, it was—like the Roman system—based on nominalism: records of debts were kept in the money of account, and coins were valued in that money of account. Over the decades and centuries following Charlemagne, debasement, clipping, and deterioration all reduced the silver content of the actual coins—to only 3 percent of the original. (Ingham 2004a, p. 189) However, prices were in the money of account—not determined by the quantity of metal in the coins.

Sometimes powerful kingdoms could coin money and reestablish the relation between actual coins and the money of account. (Ingham 2004a, p. 190) During the 12–13th centuries, more powerful political jurisdictions asserted sovereignty and claimed their own moneys of account—usually variations on the Carolingian pounds—and could impose exchange values on foreign coins that circulated across imprecise political boundaries. (ibid., p. 190–1) Otherwise, coins had different values in different zones of sovereignty—confirming that value was not determined by metal content—with exchange rates determined in the sovereign’s nominal money of account. As Ingham puts it, coins were *counted*, not *assayed* to determine metal content. (Ingham 2004a, p. 191)

Banks were operating again in Western Europe as early as the eleventh and twelfth centuries. There were four main types of banking institutions in the early Middle Ages: pawnbrokers, money changers, deposit bankers, and merchant bankers (the elite). (Lopez 1979, pp. 6–7) As commerce grew, new opportunities emerged for the merchant banks to grant credit based on trust. (op. cit., p. 9) Early deposit banks in Italy created *giro* accounts, which could be used to settle payments where both the payor and payee had accounts with the same bank (Cameron 1967, p. 315). The government of Genoa began to regulate banks in the 12th century, requiring them to keep cash and record books available for inspection; in return, the government guaranteed the credibility of the bankers and recognized entries in their books as legal proof of transactions. (Lopez 1979, p. 11)

In the 13th century, the West “rediscovered bills of exchange” (Braudel 1973, p. 359), and these “financed the bulk of the merchandise trade between the Mediterranean and northern Europe in the second half of the 13th century.” (Day 1987, p. 173) Proscriptions against usury played a role in financial developments, including the rising importance of bills of exchange. In 1049, the Council of Reims had proclaimed “no cleric or layman should be a usurer.”

(Lopez 1979, p. 4) Laws against usury viewed it “specifically and exclusively as a fixed, predetermined, certain return on a loan.” (Munro 1979, p. 170) Bills of exchange could conceal the rate of interest within the exchange rate on bills. Because the exchange rates between currencies could fluctuate, the return on these instruments was permitted as a reward for risk. However, the prohibition against usury slowed the development of the negotiable bill of exchange (discussed below) because open discounting was prohibited. The laws against usury, however, were not the principal reason for the development of the bills of exchange, for other methods existed that would permit the concealment of interest. For example, interest charges could be deducted in advance from loans; late payment charges could be added (with the late payment prearranged); or two prices could be quoted, with a higher one for deferred payment.

Development of Banking in Europe in the Pre-Modern Period

Money changing, deposit banking, and credit instruments grew together in late medieval Europe. According to Ingham, “By the late 15th century, Pacioli, in his famous treatise on double-entry bookkeeping, listed nine ways by which payment could be made,” including cash, credit, bills of exchange, and assignment in a bank. (Ingham 2004a, p. 192) Banks used overdrafts to expand credit or created *fictitious* bills of exchange: a merchant or banker would charge a banker with supplying foreign exchange in a foreign place but would agree to waive repayment abroad in order to receive postponed payment in the currency and place of origin. (Lopez 1979, p. 15)

We can look at the development of banking in Italy and Aragon as examples of the evolution of banking in Europe during the premodern period. In Italy, banking developed out of money changing, with banks operating in Lucca by the twelfth and thirteenth centuries. They changed petty foreign coins into legal Lucchese tender, lent sums in foreign coins to clients, dealt in gold and silver (accepted raw gold, sold it to gold-beaters and bought thread and leaves for resale), made loans, and accepted deposits from clients (both time and demand deposits). (Blomquist 1979, p. 60) Depositors were usually from the middle and upper classes, while loans were often made to peasants who borrowed coins to buy seeds or tools and repaid the loans in kind with grain or wine. (op. cit., pp. 63–4) The typical loan was smaller than the typical deposit, but deposits were generally longer term than loans.

By the mid-13th century, money changers made frequent and substantial loans to merchants, while as early as 1200, a money changer would settle a debt for one client by transferring the amount owed to the account of the creditor. (Blomquist 1979, pp. 65–7) Money changers began to keep accounts with one another “to facilitate settlement of debts through bank transfer.” (op. cit., p. 67) Loans were sometimes allowed by providing overdrafts to depositors.

Thus, we observe three banking functions performed by money changers as early as the 13th century: the *giro* function, the clearing mechanism, and credit creation. By the 14th century, banks were operating in the major port towns. These offered *giro* accounts used to settle accounts between customers and acted as intermediaries, loaning deposits. They also created *banco* money (notes) that would circulate in the community and expand the volume of the bank's business. (Day 1987, p. 2) Italian money changers and deposit banks operated primarily in the local market, while international commerce was in the hands of mercantile banking partnerships operating in the major towns by the mid-13th century. (Blomquist 1979, p. 68)

International banking developed as early as the 12th century in Lucca; by the 1150s, Lucca was involved in the six fairs held annually in Champagne—sending silk and buying products to bring home. This led to the development of an organized money market in Lucca: the buyer of exchange “delivered funds in Lucca and received from the seller, the taker, a notarial instrument promising repayment at one of the Champagne fairs in an equivalent amount of money of Provins.” (Blomquist 1979, p. 71) Since repayment would occur in the future, the buyer of exchange was a lender and the seller a borrower, “with interest on the buyer's capital built into the fluctuating rates of exchange.” (ibid.) For example, the borrower (seller) would receive more local (*Lucchese*) currency per foreign (*Proviniois*) currency to be delivered at the fair the closer the borrowing date was to the date of the fair. (op. cit., p. 74)

Detailed records of the Medici bank of Florence from 1427 allow fairly accurate estimation of the importance of finance during this period in one of Italy's city-states. Loans made up about a third of the bank's assets, while accounts payable equaled about sixty percent of liabilities. The major financial instruments were coins, bills of exchange, trade credit, and government debt (which equaled one-third of total financial instruments). The aggregate quantity of financial assets in Florence was equal to one-half of the total value of real assets. However, financial institution liabilities were equal to only about five percent of total assets. In other words, there was a large financial system, but most of it remained outside the banks. (Goldsmith 1987, pp. 159–169)

Banking in late medieval and early modern Aragon displays many modern characteristics, including substantial government participation in finance. Jewish banking was well established by the 12th century, and Christian banking was developed by the thirteenth. Interest charges were authorized in 1217 (although they had existed unofficially for generations), and maximum interest rates were set at 20 percent in 1229. (Riu 1979, p. 137) In 1235, the crown ruled that Christian bankers could lend at 12 percent, while Jewish bankers could lend at 20 percent. (ibid., p. 138) In 1284, the books of the money changer/banker were accepted as proof that transfers had occurred; at the same time, various regulations were placed on bankers (for example, they had

to insure their banks with the government so that the funds of depositors were guaranteed).

The first laws concerning fraud and bankruptcy were passed around 1300, after the first bank failures occurred in the late 13th century. The city of Barcelona began to use private banks as a treasury, depositing funds and relying on banks to float city debt (at an interest rate of 20 percent). (Riu 1979, p. 143) Due to continuing insecurity and the failure of private banks, Catalonia founded a public municipal bank at the beginning of the 15th century, and Valencia and Palma de Majorca soon followed suit. (op. cit., p. 149) The municipal banks operated as a treasury for the cities, offered saving accounts to private citizens, made loans to cities (typically at 5 percent interest), and offered demand deposits with checking privileges to private citizens, communities, and public corporations. Certain agents were required to hold deposits in the municipal banks, such as executors of wills, guardians, and court officials. These banks held cash against the deposits of private citizens, but deposits of businesses and others were balanced against loans to the state, official organizations, and private individuals. (op. cit., p. 151) Private banks held reserves in the municipal banks and borrowed from them to meet payments. Money orders and checks were used by 1436 to transfer funds between accounts. Between 1441 and 1450, laws were passed to decrease private competition with the municipal banks. (op. cit., p. 153) For example, private banks were forbidden to offer credit based on bills of exchange.

The public banks began the process of merging deposit banking and finance of city-states, especially for war expenses in Barcelona in 1401, Genoa in 1407, and in Venice's Bank of Rialto 1587. (Ingham 2004a, pp. 194–5) Banks would substitute their promise to pay for the state's promise to repay the public bank. Suppliers

to the state could draw on the account to make their own payments by bank transfer ... This process depended on the trustworthiness of the banks, which, in turn, relied on the legitimacy and viability of the state ... Politics and money were becoming entwined in new and more intimate ways that were eventually to be expressed in the kind of constitutional settlements that made possible the formation of institutions like the Bank of England in 1694.” (Ingham 2004a, p. 195)

However, bank failures and liquidity problems made banking fragile, especially before the creation of widespread transferability of debt:

... until the bills became transferable as a means of payment to third parties outside the network of bankers it remained “private” money. ... During the sixteenth century, bills began to leak out of the network of exchange bankers’ but remained a restricted means of payment. They were opposed in nations with strong bankers because they would lose the profits. (Ingham 2004a, p. 200)

Bills of exchange circulated upon endorsement as early as 1410 (Braudel 1973, p. 359). In 1437, the London Mayor's Court issued a decision that "recognized the transferability of a formal bill of exchange and the bearer's legal claim to full payment." (Munro 1979, p. 215) However, full discounting wasn't recognized until the 18th century, although some evidence exists of discounting of *letters obligatory* at the Antwerp fairs as early as 1536. (ibid.) Goldsmiths issued notes against deposited gold in the mid-17th century (by 1666 goldsmiths had issued 1.2 million pounds sterling). In Venice, banks had issued redeemable notes since the 15th century, and bank notes were commonly issued by the 1660s. (Braudel 1973, p. 360)

This introduction to ancient and premodern financial institutions and to the evolution of banking gives some idea of the range of practices and instruments available before the development of a monetary system with most production for the market and with most individuals involved significantly in market activities. Credit in such societies was primarily created to finance trade and had less to do with production. In such pre-capitalist societies, money and credit transferred purchasing power across time but served primarily to "lubricate" trade. As we see next, money and credit become much more important as capitalism develops. The example of Aragon, however, serves as an introduction to the increasing involvement of the state in financial practices, a subject to which we now turn.

FINANCIAL DEVELOPMENTS UP TO THE AGE OF BULLIONISM (AND CONQUEST)

In the societies previously examined, credit was commonly used, but these were not economies dominated by production for the market. As discussed in Chapter 2, the state played a major role in creating markets, and prices were set by treaties negotiated by states. Further, as discussed above, the medieval city-state also played a role as the guarantor and enforcer of contracts. However, during the age of conquest of the New World, the state became much more directly involved in the creation of monetary instruments and institutions, primarily to finance military adventures, although the high consumption levels of the crown cannot be ignored. We can define this as the age of bullionism—in which wars were fought by and for precious metal. At times, some nations tried to maintain a strict relation between the nominal value of a coin and the embodied precious metal—although "debasement" was common and nominalism usually was adopted. Let's begin by looking at early government finance.

Italian cities had a long tradition (dating back to Roman times) of "farming out" tax collection to individuals who would retain a portion of the revenues. In a time of unusual need for revenue, the city could "anticipate" it by

selling the right to collect future taxes. As early as 1164, the Venetian Republic obtained an advance on future tax revenues (Ehrenberg n.d., p. 47). In Venice, the Bank of St. George monetized the city's debt as early as the mid-15th century, and the *Banco del Giro* was set up in 1619 specifically to provide finance to the government. (ibid.) Even as late as the 16th century, cities expressly gave creditors the right to the person and property of all citizens should the city default on debt, making city debt relatively safe. (ibid., p. 33) Indeed, in Genoa, the debts of the city were so liquid that they circulated as freely as currency, functioning as a means of payment. (ibid., p. 48)

While republics generally were able to borrow, monarchies were typically more limited in their ability to obtain loans—especially to finance war. While the crown could also anticipate revenues, its subjects were not normally liable for crown debts. In addition to advances on future tax revenues, the crown could resort to the sale of offices, land sales, net bullion inflows, and debasement of coins. Of these, only the last two could be a major continuing source of revenue. The policy of *Bullionism* was designed to increase inflows for coinage. While debasement and recoinage could conserve bullion, the state mint was limited in its ability to provide purchasing power through debasement because traders might reject debased coins. Furthermore, it must be remembered that wars would be fought largely on foreign ground, using local supplies and hired mercenaries. Typically, traders in foreign lands based the value of a coin on its metallic content and not on its nominal face value due to uncertainty about the war's outcome. (Ehrenberg n.d., p. 31) Essentially wars on foreign ground were fought with—and for—gold and silver. With the “discovery” of the New World at the end of the 15th century, another source of bullion—and of war—was created.

As the conduct of war passed from the hands of the citizenry to hired mercenaries, the ability to generate revenue became increasingly important. During the thirteenth and fourteenth centuries, Italian cities increasingly hired the *Condottieri* (professional private undertakers) for military purposes. Given the ability of the city republics to obtain advances on revenues, financing of the mercenaries could normally be met through the expansion of credit. In general, military spending accounted for one-third to two-thirds of total government spending. In Medici Florence, military spending alone during the period 1421–30 was twice total government revenues (Goldsmith 1987, pp. 164 and 249). The Genoese debt in 1408 “was equal to about four times the total volume of Genoese maritime trade that year,” with much of that debt accumulated during a war with Venice. (Day 1987, p. 158) As another example, “the expenditure of the Spanish Crown in putting down the rebellion in the Netherlands averaged two to three million gold crowns a year, i.e. more than the yearly revenue of the Netherlands Government during the most flourishing trade period.” (Ehrenberg n.d., p. 28) Furthermore, “in the sixteenth century there were only

twenty-five years, in the seventeenth century only twenty-one years, in which there were no war-like operations on a large scale.” (op. cit., p. 28) After 1585, the Elizabethan government’s spending on the military accounted for four-fifths of total government spending. (Goldsmith 1987, p. 193) Thus, the enormous costs of war became a normal expense that had to be met.

While city republics could meet this through credit creation, monarchies had more difficulty. It was impossible for a king to build a war chest large enough to conduct even a small war without borrowing. The result was a never-ending cycle of debt. Typically, the crown could only borrow on the anticipation of a specified revenue source. Even then, the crown frequently had to obtain a guarantee from a group of respected wealthy individuals, from the church, or from a city to obtain an advance, normally made only at high interest rates—even as high as 42 percent. (Prestwich 1979, p. 87) However, the high interest rates and the continuing needs of war finance could force the crown to default on interest payments and even on principal. (Ehrenberg n.d., pp. 39–41) Indeed, sanctions against usury frequently arose because the crown found it expedient to abrogate its own interest commitments. (op. cit., pp. 34, 43) For example, in England, interest charges on loans made to the crown were explicit until 1240, when Henry III forbade usury to reduce his own interest payments. (Prestwich 1979, p. 85)

If interest and principal were often unpaid, why would anyone lend to the crown? In some cases, the loans were forced, but many were made voluntarily in return for special treatment. For example, Italian lenders to the English crown received in return for the loans that were rarely repaid: land; exemptions from taxation and jury service; commercial advantages; use of the exchequer’s machinery against clients who defaulted on their obligations; and favorable treatment in courts. (Prestwich 1979, p. 91) In spite of these concessions, even by the end of the Middle Ages, most of the crowns of Europe were hindered in their ability to raise revenue through loans.

There were three primary means through which the state could increase its purchasing power. First, it could make its subjects responsible for state debts—as Italian city-states did. Citizenship, patriotism, and representative government all contributed to this development, but are beyond our scope. Second, the state could increase its ability to generate revenue through Bullionism: a policy to increase the flow of precious metal to the crown. Finally, the state could develop fiat money. This would require, however, that state money become at least as desirable as private credit money—otherwise, state money would circulate at a discount within the *giro*. Let’s examine Bullionism and the creation of fiat currency, in turn.

Bullionism was encouraged in the 16th century by the crown’s chronic shortfall of revenues and by its need to conduct foreign wars. This partially explains the concern with finding new sources of gold and silver, and the resulting race

to colonize the Americas. As Columbus explains in his ship's log from the first voyage to the New World, the explicit purpose was to obtain gold to fund "the conquest of the Holy Land. I have already petitioned Your Highnesses to see that all the profits of this, my enterprise, should be spent on the conquest of Jerusalem, and Your Highnesses smiled and said that the idea pleased them, and that even without this expedition they had the inclination to do it." (Fuson 1992, p. 157)

Alternatively, if a trade surplus could be generated that would lead to an inflow of coined precious metals, then the crown could tax a portion of that inflow to meet its military and other payments. Above all, the crown had to prevent the export of precious metal. However, trade was typically based on bills of exchange. For example, the English wool export trade required payment by *drapers*, the purchasers, of one-third in cash, while the remainder was received in the form of two bills: one payable in six months and the other in twelve. As early as 1340, the English crown tried to force *staplers* (wool exporters) to pay a tax in the form of precious metal coins. The *staplers* responded that Flanders (importer of wool) forbade the export of bullion so that payment of the tax was impossible, and the tax was repealed. The crown tried again in 1429, requiring them to demand full payment in the form of English coin, and to bring one-third to the English mint. However, the *staplers* complained that the ban on the use of credit was destroying the export market, and the law was finally repealed in the 1470s.

BULLIONISM, STATE FIAT MONEY, AND THE GIRO

The state in the late premodern period was outside the *giro*. It needed a source of spending power to conduct its wars, but because its own debts were not always acceptable, it was forced to rely on bullion (to coin) for payment—even if the value of coins could be maintained above the value of the embodied metal, bullionism keeps the state outside the *giro*. Private traders are thereby encouraged to continue to use privately created credit money—especially if they can remain outside the crown's taxing power. As long as trade can occur based on private *giro* money (such as bills of exchange), the state's ability to raise revenues remains limited.

The state could respond by debasing the coin—reducing the embodied metal. In 1250, the English pound sterling contained 324 grams of pure silver, but this had fallen to 112 grams by 1600. Similarly, between the years 800 and 1600, the French monetary unit fell from 390 grams to 11 grams, the Milanese unit fell from 390 grams to 4.9 grams, and the Venetian unit fell from 390 grams to 3.5 grams. (Cipolla 1980, p. 201)

With the state outside the *giro*, its subjects might object to debasement. Traders could turn to their private *giro*—with credit denominated in a private

money of account. For example, in 1619, the Hamburg *giro* bank developed a unit called the *mark banco* that was used for transactions within the private *giro*. This unit fluctuated in value against the state monetary unit (*thaler*). (Knapp 1924, pp. 144–9) This is also cited by Day, who reported the “Genoese created a perfectly stable accounting unit or *moneta di fiera* (fair money), the *scudo di marco ...*” of constant value against gold. (Day 1987, p. 148)

Still, courts could rule in favor of nominalism—forcing creditors to accept payment in coins at current announced value. For example, in 13th century Britain,

The evidence of a range of relevant legal sources strongly suggests that, from at least the late thirteenth century, the common law took what would now a days be called a “nominalist” approach to the valuation of money. Monetary obligations were expected to be paid at their money of account value using coins which were current at the date of payment. Any change to the monetary standard between the dates of contract and payment was generally irrelevant to the performance of the obligation. It seems that the common law courts would not revalue the obligor’s debt to allow for any adjustment to the monetary standard during that time. (Fox 2016, p. 203)

Fox goes on:

It will be remembered that the author of the *Tractatus Nove Monete* (c.1286–7) said that any new issue of coins was made “generally known by public proclamation in the accustomed way by the prince’s crier”. (op. cit., p. 208)

The final point of the quotation from the *Tractatus Nove Monete* was that the sovereign’s money must ‘not be refused by any of the public without penalty’. The public owed a duty directly to the sovereign to tender and accept his or her money according to the legal valuation ... [and King Henry VIII’s Proclamation stated the consequences of refusing the crown’s coin was] “the same person or persons so refusing or denying, and to commit him or them to ward and prison, there to remain without bail or mainprize unto such time as the King’s determinate pleasure be further known in that behalf.” (op. cit., p. 209)

Note that nominalism did not apply to foreign currency.

The English courts of the late medieval and early modern periods were familiar with enforcing English transactions denominated in foreign currencies. But unlike the English sovereign’s coins, foreign moneys did not have a value which a judge could—let alone must—recognize. Their value in terms of English currency was always a question of fact for the jury. Indeed, it would eventually be said that foreign money was no different from bullion in legal estimation. This was not to say that a debt to pay foreign money could be discharged by tendering uncoined bullion (*massa*). The point was that both bullion and foreign coin had to be valued by evidence presented to a jury. (op. cit., p. 207)

Only a state using fiat money whose value is set in the money of account, can gain purchasing power by issuing currency. This was recognized by Keynes, who argued:

When ... a coin is no more than a quantity of bullion, of which the stamp may certify the quality and indicate the quantity, [it] ... will not circulate except for its bullion value. In this elementary stage the expedient of debasement is not available. It cannot appear, until with the development of contract the conception of a money of account has emerged, and the coins issued by a state have acquired the character of legal tender and enjoy a *cours force* as the legal discharge of obligations calculated in this money of account. It is at this stage that money, in the sense in which we understand it, makes its entry into human institutions. (Keynes 1982, p. 226)

The modern state can become a member of the *giro* by accepting the liabilities of banks in payment of taxes and in other dealings with the state. When the state becomes a part of the *giro*, the liabilities of banks become the generally accepted means of payment. At this point, the banknote is not only useful in dealings with the bank and with the bank's customers, but also in dealings with the state, now a member of the *giro*. This expands the state's ability to spend. It also makes the banknote acceptable to those who were formerly outside the bank *giro* since those with obligations to the state become members of the *giro* by virtue of their relations to the state. The entrance of the state enhances the *giro* even as it enhances the financial strength of the state.

The state's power can be increased if it rises to the apex of the *giro*. To do this, it must ensure that its liabilities are preferred above those of banks. As discussed below, this is a long process that accompanies the development of a central bank and a mono-reserve system. The state gives impetus to the expansion of bank business by accepting bank liabilities, but also encourages convertibility of bank notes (redemption of bank liabilities for government liabilities). Convertibility of bank liabilities is not most importantly a method used to control the issue of notes (as in the deposit multiplier notion), but of ensuring the value of state money. The state also restricts banks by eventually eliminating private notes so that only central bank notes will act as currency. This facilitates the movement to a system based on demand deposits against which reserves are held at the central bank.

Thus, the state is gradually transformed from a relatively impoverished borrower that must rely on private guarantees of its credit worthiness to the modern capitalist state whose debts are the most preferred and whose guarantees lend credit worthiness to private debt issues. Again, this transition is partially based on the creation of patriotism, the development of republican forms of government, and other (primarily) non-economic matters that are beyond our scope. However, a brief study of the development of banking and the central bank in England will show how resolving the chronic fiscal problems of the

pre-modern state also helped to bring on modern banking. That is, the rise of the modern state and the rise of the modern financial system are inextricably related. Indeed, it would not be far from the truth to argue that the state forced modernization onto the monetary system. (Heilbroner 1985, p. 88).

THE DEVELOPMENT OF CENTRAL BANKING AND STRENGTHENING OF THE SOVEREIGN'S CURRENCY: THE CASE OF ENGLAND

In this section, we will examine British monetary history from the 14th through the 18th century. In the early years, Italian bankers were the primary lenders to the crown. The crown also issued metallic coins as well as tally sticks. (Maddox 1969) Private commerce took place in the sovereign's coin, bills of exchange, and other forms of credit. Private banks issued notes, with country banks using London banks for clearing. Two key events led to a fundamental reformation of the financial system that gradually led to the modern banking arrangements that we still have today: (1) a default by the King on tally stick debts led to the creation of the Bank of England, which became the nation's central bank, and (2) Queen Elizabeth's monetary reformation that somewhat paradoxically established a strong link between the currency and precious metal. The first would eventually bring the state into the banking *giro* system, and while the second would seem to be more consistent with Bullionism, it ultimately strengthened the position of the pound sterling—a monetary system fundamentally based on nominalism.

The Transformation of State Finance: Italian Bankers, Monetary Reform, and the Creation of the Bank of England

While currency exchange played a major role in the development of the Italian financial system, state finance shaped the development of the English financial system. Italians dominated English finance in the late thirteenth and early fourteenth centuries and were primarily involved in the export of wool, loans to the crown, and payments made on the Continent for the English. Loans to the crown enabled the nation to conduct foreign wars (Italians loaned over £125,000 sterling in 1338–39 at the outbreak of the Hundred Years War). (Prestwich 1979, p. 79) The loans were made at high interest rates and were rarely repaid in full. However, the Italians received special benefits in return for their services. Italian bankers did not play a big role as depositories but did accept short-term funds to make payments on behalf of the depositors. (op. cit., p. 96)

Following the principle of Bullionism, the crown had strict prohibitions on the export of coin, so Italian bankers would use bills of exchange, accepting

pounds in England, then delivering, for example, *livre tournois* in Paris. (ibid., p. 97) The crown also maintained strict control over exchanging foreign coins for sterling: it was permitted only at the royal exchange. This was to prevent the export of English bullion and to prohibit the import of counterfeit and clipped coin. These controls meant that domestic banking couldn't develop out of currency exchange in England, as it had in Lucca. (Prestwich 1979, p. 99)

Detailed evidence shows that small credits played an essential role in pre-modern England before the development of an extensive banking system. (Mcintosh 1988, p. 560) Consumers would settle accounts with coins, but merchants would use them only to settle net debts. The evidence shows that debts might be carried for months (and even for years) on the accounts of merchants (and consumers) before finally being settled.

Any two people might build up a number of outstanding debts to each other. As long as goodwill between the individuals remained firm, the balances could go uncollected for years. When the parties chose to settle on an amicable basis, they normally named auditors who totaled all current unpaid debts or deliveries and determined the sum which had to be paid to clear the slate. If trust between the parties broke down, the complainant could bring suit in the Havering court ... (Mcintosh 1988, p. 561)

Such suits were common.

Still, in many ways, England's financial system was repressed as late as the 17th century compared to Europe's. By that time, two forms of money had spread across Europe: private credit instruments and public metal coins. As Ingham (2004a) argues, in England, these were in uneasy conflict. The sovereign power of English kings was, in a sense, too great—they were immune from legal action for default and not responsible for their predecessor's debts, so lenders were reluctant to take their debt. Thus, they relied largely on coin, plus borrowing from Italian bankers. (See also Ingham 2004b, Chapter 6.)

They typically practiced Bullionism, and “from the fourteenth to the mid-17th century, [they] banned the importation of foreign coins and the export of bullion; commanded exporters to supply their bullion to the mints; attempted to prohibit the bill of exchange; and generally sought to limit the use of credit.” (Ingham 2004a, p. 203) In 1560–1, Queen Elizabeth I reformed coins, adopting four ounces of silver as the standard for the pound, a standard that lasted until WWI. It was the “lynchpin of England's fiscal and political system. Its retention was a condition of the survival of the constitutional settlement between sovereign, government and ruling classes ...” and “encouraged a steady supply of long-term creditors for the state and in this way provided a secure basis for the eventual adoption and expansion of the credit money system.” (Ingham 2004a, p. 204) Her reform established a coinage and prohibited foreign coins, so the domestic money was coextensive with the state and “English monetary

policy was unequivocally monarchical and bullionist.” The “emerging English nation state became the basis for the impersonal trust that eventually enabled the forms of credit money to become established outside interpersonal banking and exchange networks.” (ibid.)

Ironically, then, it was the backwardness that helped England to surge ahead to create the modern financial system that was needed for the development of capitalism, in which the state and its central bank establish the money of account and issue the currency, the reserves used by private banks, and the safe government debt that underlies the financial system. Ingham (2004a) explains how the strong links of the currency to precious metals, the transformation of bills of exchange into impersonal and transferable debt, and the creation of the Bank of England fostered the evolution to the modern financial system. (See also Ingham 2013a.)

One final apparent setback played a key role in developing the central bank. King Charles II had borrowed against expected revenue, issuing tally stick receipts for loans. With war against France coming and the crown deeply in debt, he defaulted on the tallies in 1672. London money interests were so upset that they invited invasion and seizure of the throne by William of Orange. The constitutional settlement of 1689 provided the new king with too little income so that he would have to rely on Parliament for funding—putting the power of the purse into Parliament’s hands. (Ingham 2004a, p. 208–9; See also Desan 2014.)

In 1694, backed by creditors of the King, Parliament created the Bank of England to provide loans to a Whig government at war with France and in dire financial straits. The Bank of England was incorporated as a private bank with special privileges: it was created as the sole depository for state balances; it was the only firm incorporated with limited liability for stockholders; in 1697, Parliament declared that no other bank would receive a state charter; and in 1708, Parliament rewarded the Bank of England for taking up government bonds by prohibiting any other joint stock bank with more than six partners from issuing notes in England. (Bagehot 1927, pp. 90–95; White 1984, p. 38)

At its creation, the Bank of England issued stock, using the revenue raised to make a loan to the King. The loan was to be repaid by future customs and excise taxes. The Bank would issue notes (supposedly backed by the King’s debt) and lend those—earning interest from the King and from the borrowers! That converted the King’s personal debt into a public debt and the notes into a public currency. Ingham argues: “Underpinning this transformation in the social production of money was the change in the balance of power expressed in the ‘hybridised’ concept of sovereignty of the ‘king-in-parliament’.” (Ingham 2004a, p. 209)

Private Banking in England and Development of Central Banking

Until 1826, there were three other categories of English banks: private London banks that issued notes; private note-issuing country banks; and joint stock banks that could not issue notes. The special status the Bank of England enjoyed as a note-issuing joint stock bank gave it a tremendous advantage: as a joint stock bank, it could raise a large sum of capital (1.2 million pounds) through subscription. In contrast, country banks that issued notes were chronically undercapitalized because they were limited to six partners (White 1984, p. 38). Furthermore, as a limited liability firm, the Bank of England's owners could protect their fortunes—unlike private bankers and the owners of other joint stock banks, who placed their entire fortunes on the line.

Banks did not begin as depositories of funds. Early banks engaged in three activities: providing loans to the government, changing coins, and facilitating the circulation of bills of exchange. (Bagehot 1927, p. 77) Gradually, however, banks moved from simply helping to circulate bills of exchange to discounting them by providing bank notes. As Bagehot put it, “[N]o nation as yet has arrived at a great system of deposit banking without going first through the preliminary stage of note issue ... ” (ibid., p. 88) As late as 1800, bank notes accounted for 50 percent of M1 in England, while deposits accounted for only 10 percent. (Cameron 1967, p. 8) Deposit banking dominated only later, after the public had become accustomed to using bank liabilities as media of exchange, means of payment, and stores of value.

Thus, by creating this special right of note issue, Parliament provided an advantage that would help the Bank of England to monopolize note issuance in London. Although private banks could legally issue notes in the city, their notes had been almost completely displaced by Bank of England notes by the mid-18th century. (Bagehot 1927, p. 96) Since checkable deposit banking was not yet important, this meant the Bank of England “became the bank in London.” (ibid., p. 97) Other London banks focused on foreign trade, relations with stock and bill brokers, and relations with the country banks.

With the creation of the Bank of England, the “emerging English nation state became the basis for the impersonal trust that eventually enabled the forms of credit money to become established outside interpersonal banking and exchange networks.” (Ingham 2004a, p. 205) The Bank's monopoly over discounting bills of exchange and note issue gradually integrated the banking system under the Bank's discount rate. At the same time, it linked private money and public money as well as public debt and private debt. (Ingham 2004a, pp. 210–11) Bills of exchange were gradually delinked from particular commodities (called “dry exchange”) and transformed into a pure form of credit; they were also gradually detached from particular creditors and debtors to become transferable. In short, they became increasingly “depersonalized”

(payable to bearer) and issued as bank money (in the form of notes). (Ingham 2004a, pp. 187, 199)

This required a particular social and political structure, with the Crown ceding power to Parliament. The Crown's IOUs became "national debts" (like Italian city-state debts) and the basis for public credit money. A genuinely impersonal sphere of exchange was eventually provided by the nation state. (Ingham 2004a, p. 202)

As the largest makers and receivers of payments and in declaring what was acceptable as a payment of taxes, states were the ultimate arbiters of currency. They created monetary spaces that integrated social groups whose interaction was not embedded in particular social ties or specific economic interests. Until credit money was incorporated into the fiscal system of states which commanded a secure jurisdiction involving extensive legitimacy, it remained, in evolutionary terms a 'dead-end'. (ibid.)

One might think that all these developments would have eliminated Bullionism in England: the central bank could issue notes and lend to the crown. Surely precious metal was no longer needed? Instead of dropping precious metal, the English commitment was paradoxically strengthened: by 1727 the pound was firmly backed by silver (although a bi-metallic standard was often maintained—setting both the price of gold and silver—and a gold standard was adopted in the 19th century). Credit money had become the most common means of payment, but England also had the strongest metallic currency in history. (Ingham 2004a, p. 210–11) Ingham argues that this strengthened the English moneys (coins, notes, and bills) through formal convertibility of the pound to precious metal. It was not necessary to maintain a constant amount of precious metals in the coins because it was the pound—not the coin—that was fixed to silver; the value of coins was fixed to the pound. That convertibility (of coins to pounds and pounds to silver) reduced the fear that the sovereign would "cry down" the value of the coin (require more coins to make up a pound sterling—remember that the pound, itself, was never coined), or default altogether, refusing the sovereign's own debts in payment.

The Bank Restriction and Rising Dominance of Bank of England Notes

Remarkably, however, between 1797 and 1819, the Bank of England was prohibited from redeeming its notes for bullion. The "Bank Restriction," as it was called, ensured that a run on the Bank of England could not occur. "A bank of issue, which need not pay its notes in cash has a charmed life; it can lend what it wishes, and issue what it likes, with no fear of harm to itself, and with no substantial check but its own inclination." (Bagehot 1927, p. 107) (Of course, Bagehot exaggerates since note issue is always constrained by the willingness

of the public to accept notes.) Furthermore, “since 1797 the public have always expected the Government to help the Bank if necessary.” (ibid., p. 108) Thus, the Restriction strengthened public confidence in the Bank rather than reducing it as one might expect suspension of convertibility to do.

Bank of England notes were accepted by the government in payment of taxes and customs duties and for subscriptions to government loans. (Cameron 1967, p. 22) This helped to increase the desirability of the notes. Parliament also gave the Bank’s notes special status, allowing country banks to make their notes convertible into its notes in 1833. (White 1984, p. 39) In this way, Bank of England notes would become as desirable as coined money since they would replace coins as the banking system reserve. Later, mere deposits at the Bank would serve as the reserve for London banks, while deposits at London banks would serve as the reserve for country banks.

To increase the circulation of Bank of England notes outside London, Parliament periodically restricted the note issue by country banks. For example, Parliament prohibited any bank other than the Bank of England from issuing one-pound notes after 1775, and the five-pound note was banned in 1777. These restrictions were later temporarily dropped, but the prohibition on the one-pound note was restored in 1829. (Wood 1939, p. 38) In 1826, Parliament encouraged the Bank of England to open branches in cities outside London to increase the circulation of its notes. (op. cit., p. 39) By 1844, the quantity of Bank of England notes circulating in England had risen to equality with that of private notes. (op. cit., p. 23)

The creation of the Bank of England thus consolidated the public and private monetary systems: “This fusion of the two monies, which England’s political settlement and rejection of absolutist monetary sovereignty had made possible” solved two problems: the private money of the bill of exchange became more widely circulated based on impersonal trust and legitimacy, and Parliament sanctioned tax and duty collection to service interest on sovereign loans. (Ingham 2004a, p. 209–10) England struck a balance between too much and too little sovereign power—the sovereign was no longer a credit risk, and as a result, the sovereign could operate with higher debts and taxes. The “emerging English nation state became the basis for the impersonal trust that eventually enabled the forms of credit money to become established outside interpersonal banking and exchange networks.” (Ingham 2004a, p. 205)

THE CENTRAL BANK, RESERVES, AND DEVELOPMENT OF MODERN BANKING

In this section, we will look at the development of central banking and its relation to private banks.

Pyramiding Reserves on London

London had been the center of English foreign trade since medieval times, with the bill of exchange drawn on London as the primary international means of payment. Taxes had to be paid in London. London also acted as the clearing house for country bankers (although country bankers in a town or small region would operate a local clearing exchange that would not require intervention by London), thus they typically held deposits as a primary reserve in a London city bank. These could be used for clearing or if the need for cash arose. The country bank would also leave stocks and bonds in London as secondary reserves, against which overdrafts could be drawn. (Sayers 1957, p. 109) The quantity of reserves varied greatly and was frequently overdrawn. It was not uncommon for a country bank to have an overdraft outstanding equal to fully 10 percent of the bank's total balance sheet. (op. cit., p. 121)

Thus, London banks became the depositories for country bank primary and secondary reserves because London was the financial center, and London banks frequently paid 4 percent interest on country bank deposits, charging 5 percent on overdrafts. The normal deposit of reserves required of country banks was £10,000—it was not determined by the size of the country bank's balance sheet. (Sayers 1957, p. 35) If country banks wished to hold more 'liquidity', they would send more stocks and bonds to their correspondent to increase secondary reserves. (Wood 1939, p. 21) London banks would use the primary reserves to earn interest and overdrafts as a way to influence the behavior of country banks. (Sayers 1957, p. 124) Country banks issued notes payable on accounts in correspondent London banks. In this way, the clearing mechanism was enhanced, and the acceptability of notes was improved because country bank notes could then freely circulate outside their limited *giro*, since other country banks could send the notes to London for clearing.

In turn, London banks held Bank of England notes as their reserve. After 1833, as noted above, country bank notes were convertible into Bank of England notes, which made the liabilities of the Bank of England even more desirable. Gradually, London banks (and even later, country banks) substituted deposits at the Bank of England for notes held as reserves. (Wood 1939, p. 178) Thus, England developed a mono-reserve system: all reserves ultimately were supported by notes or deposits issued by the Bank of England—the country banks held reserves at London banks, the brokers held reserves at London banks or at the Bank of England, and the London banks held reserves at the Bank of England. "No other bank holds any amount of substantial importance in its own till beyond what is wanted for daily purposes." (Bagehot 1927, p. 28)

The Act of 1844 separated the Bank of England into two departments: the Note Issue Department and the Banking Department. The Note Issue Department issued notes based on government debt and bullion held. Thus,

finance was provided to the state when the Bank issued notes against government debt. The Act originally permitted the Bank to issue £14 million of notes against securities, but this was supplemented periodically by new authorizations. The Banking Department could not issue notes but held the deposits of the state and of banks and brokers. It purchased government securities, other securities, notes, and coins, operating as a profit seeking bank. (Bagehot 1927, pp. 23–5) In 1795, government securities accounted for more than 77 percent of total assets held by the Bank of England, but by 1869, they accounted for only one-third of the assets of the Issue Department and 31 percent of the assets of the Banking Department. (Cameron 1967, p. 20, and Bagehot 1927, pp. 24–5)

Through the early 19th century, the dichotomy between country banks and London banks was strict. The Bank of England controlled note issuance in London, while country banks dominated note issuance outside London. Country bank notes were used mainly for the payment of wages and other small local payments. Until 1826, very few Bank of England notes circulated outside London, so virtually all notes in circulation in the country were those of the country banks. However, the Act of 1844 placed a ceiling on the quantity of private issues, and deposits gradually replaced notes. Furthermore, Bank of England notes began to circulate more widely in the country, reaching 50 percent of total issues by 1844. The Bank of England (wisely) accepted country bank notes (normally drawn on London), so it could remove them from circulation by presenting them for payment. (Wood 1939, pp. 17–23) In this way, the Bank of England could enhance its own circulation.

The Bank Acts, Suppression of Notes, and Sound Money Policy

After 1830 the London banks aggressively pursued relationships with country banks. Similarly, the Bank of England began to expand its activities outside London during the same period. The Acts of 1826, 1833, and 1844 gave impetus to this by centralizing note issue and by permitting unlimited joint stock banking: the Act of 1826 permitted this outside the London area, while the Act of 1833 allowed joint stock banking in London provided that the bank did not issue notes. The Act of 1844 provided for the gradual suppression of note issuance by country banks, to be fully eliminated within twelve years.

Until the 1820s, it was generally believed that the rate of growth of country bank note issuance was controlled by the Bank of England. This control was maintained through, first, convertibility of country bank notes to Bank of England notes, and second, by balance of payment flows. If country banks wanted to maintain the value of their notes, it was widely believed that they would have to agree to redeem them on demand for Bank of England notes. This was thought to require some ratio of country bank notes to Bank of

England notes. If country banks issued too many notes, the balance of payments would turn in favor of London, with the country bank notes flowing to London where they would be taken out of circulation.

However, over the course of the 1820s, it was argued that the rate of growth of country bank notes exceeded the rate of growth of Bank of England notes and led to the belief that the country banks were thwarting the Bank of England's attempt at sound monetary policy. (Sayers 1957, p. 140) A series of laws restricting bank note issues was designed to halt this practice. After 1826, the Bank of England would favor only non-issuing country banks in its discounting policy, refusing to discount paper submitted by country banks that issued bank notes. It pressured country banks to call in their notes and replace them with Bank of England notes. Many country bankers were able to resist this pressure—especially in areas where the population preferred the country bank notes. Banks could also get around restrictions on note issues by creating very short-term bills, and hostility in some areas toward the Bank of England would ensure local preference for them. (*ibid.*, pp. 151–2)

These policies initially worked in the favor of the Bank of England, and its notes displaced some country bank notes. However, the development of the bank check was probably more important in eliminating private country bank notes. While the Acts placed limits on the quantity of notes each bank could issue, actual issues quickly fell below the ceiling. (Sayers 1957, p. 151) Thus, restrictions on note issue might have had little effect on most banks.

On the other hand, the laws favoring joint stock banking had important and lasting effects. First, London banks declined in importance because the joint stock banks in the country would establish correspondence relations, so that the London banks became less important for clearing. Mergers of joint stock banks led to the development of national banking networks, which would include a bank in each major city—as well as a position on the London Clearing Bank—by the late 19th century. (Sayers 1957, p. 248)

Second, the limited liability of joint-stock banking (permitted by the Acts of 1858 and 1862) led to a movement to force such banks to publish balance sheets. (Sayers 1957, p. 224) Even by the mid-19th century, there were no established rules (or even rules of thumb) regarding the level of reserves or leverage ratios to be maintained by a prudent bank; as late as 1888, the ratio of cash plus short-term money market loans to deposits plus notes varied from 4.8 percent to 32 percent among banks. (*op. cit.*, p. 178) Indeed, this ratio would vary from year to year for an individual bank by as much, or more, than it did for banks on average. Even after published records became required, they meant very little until the 20th century since banks would engage in “window dressing” just before the publication date. However, the movement to force the publication of balance sheets would gradually make bank practices more uniform regarding such matters as reserve ratios.

Third, the evolution of joint stock banks and the centralization of decision-making led to rules of thumb and rules of behavior that also gradually established more uniform banking practices, such as the quantity of primary and secondary reserves that should be held against bank liabilities. Banks also established more sensible pricing policies. Interest rates on most country bank assets and liabilities had been set by collusion and rough rules of thumb. Even as late as the end of the 19th century, price competition was rare. (Sayers 1957, p. 160) Integration of financial markets and competition from the growing joint stock banks, however, forced country banks after 1850 to compete with London interest rates. This would gradually increase the control of the Bank of England over financial conditions throughout the nation. (*ibid.*, pp. 162–5)

Refining the Role of the Central Bank

As the single repository of reserves (and because its own notes counted as reserves), the Bank of England could cause tight money by raising the price of reserves. It could call in advances made to London banks or brokers, raise the interest rate required in its discount of bills, or simply refuse to discount bills altogether. When faced with tight money, London banks would call in overdrafts and force the country banks to sell consols or stocks held by their London correspondent. London banks would help to sell bills for their country bank clients but normally refused to “rediscount” bills themselves—it was the policy of London banks to avoid putting their own name on a bill to be sold. This tended to cause long-term interest rates to rise and the prices of bonds fell. (Sayers 1957, pp. 125–7)

The Bank of England also used open market sales to increase its own reserves and not primarily to influence credit conditions. (Wood 1939, p. 5) When the Bank of England was faced with an external drain of specie (for example, when England ran a trade deficit), it would raise interest rates to attract bullion inflows. However, an external drain would frequently lead to an internal drain: bank customers might notice a drain on bank reserves, become worried over the stability of banks, and try to obtain loans (discount bills) before credit was cut off, or they would try to withdraw deposits or redeem notes. This, of course, would lead to a further drain on the reserves at the Bank of England. A panic would result whenever the Bank of England acted like it might not provide the reserves needed by the private banks. Indeed, the Bank of England would frequently try to sell securities to replenish its own reserves during an internal drain. (Bagehot 1927, p. 65) This was because the Bank of England for quite some time saw itself not as a central bank, but as a profit-seeking bank.

When banks were small and local, a run on deposits or notes could be averted by the declaration of a local wealthy or respected family that the bank’s books

were sound. (Sayers 1957, p. 208) Local merchants could help if they agreed to continue to accept the bank's notes as a means of payment. On occasion, other banks in the region would step in to save a failing bank. As relations between London and the country banks strengthened, crises that began in London would spread to the country. If the Bank of England restricted loans, the London banks would often freeze the funds of country banks and call in overdrafts, forcing country banks to follow suit. Each financial crisis would be followed by an unwillingness among banks to accept long-term commitments and a preference for the discount of bills to overdrafts by customers. (op. cit., p. 212)

Thus, the evolution of the banking system into an integrated system of joint-stock banks, where decision-making was centralized in London, and into a mono-reserve system that placed considerable power in the hands of the Bank of England, occurred over the 19th century through the end of World War I. The Bank of England was originally founded to provide state finance and was granted special privileges as a reward. These special privileges caused its notes to circulate in London to the exclusion of all privately issued notes. Country banks established relations with London banks to facilitate note (and, later, deposit) clearing and to provide services as depositories of reserves. The special status of Bank of England notes, the special relationship of the Bank of England with the state, and legislation granting special rights to the Bank of England all enabled it to become the ultimate depository for a financial system that pyramided reserves.

This special status of the Bank of England allowed its notes to become fiat money and to eliminate the need for a rigorously maintained metallic coin, primarily because privately issued *giro* money was made convertible into Bank of England notes. This brought the state into the *giro* since state finance was provided by the Bank of England through a purchase of government securities. Because London was the center for international finance, nominal sterling became the world currency for use in international trade. Bullionism was displaced by Free Trade because precious metal was no longer needed to provide state finance, and, indeed, would hinder the growth of international trade with London as the center. The evolution of banks to a centralized financial system based in London increased the complexity of financial relations, increased leverage ratios (through increased pyramiding of reserves), made banks more similar, and led to an economy that was more integrated so that crises would quickly spread. Thus, the actions of the Bank of England would have more impact on banks far from the center, but the integration of the financial system also increased the responsibility placed on the Bank of England to prevent crises.

Bagshot correctly recognized that this responsibility would require the Bank to act against its own perceived short-run self-interest in crisis. Rather

than tightening to protect its balance sheet, it would have to act as lender of last resort. Bagehot laid out reasonable conditions for such operations that would be widely adopted by all the major central banks. The USA's own central bank, the Federal Reserve Bank, was founded in 1913 specifically to serve as a lender of last resort. Interestingly, however, it was not supposed to be involved in government finance—which was the original purpose of the Bank of England. However, just a few short years after its founding, its services were needed in WWI, and then on a much bigger scale in WWII. We will examine later how modern central banks continue to assist in financing the treasuries.

A SUMMING UP: IMPLICATIONS FOR THE TRANSITION TO CAPITALISM AND A MONETARY SYSTEM OF PRODUCTION

Let's summarize the points of the previous sections. As the economic system evolved toward the modern capitalist system, the financial system also underwent great changes. In the early premodern period, most production took place outside the market, so the monetized sector was relatively small. Goods circulated on the basis of bills of exchange (especially in long-distance trade) and coined money (in local trade), although other forms of *giro* money were also used in specific regions.

Gradually, however, private property rights were extended (including, importantly, to slaves—a topic examined later) and the importance of the market (again, including the market in slaves) increased. This required the development of sophisticated credit facilities—primarily based on note-issue by private banks. To extend the size of the *giros*, clearing facilities had to be created. This led to the development of a system that concentrated reserves to facilitate clearing.

The earliest central banks (such as the Swedish *Riksbank* and the Bank of England) were explicitly created to provide state finance. (Goodhart 1989b, p. 88) By providing special monopoly advantages, the central banks' notes became fiat money that could provide purchasing power to the state. At the same time, a mono-reserve system was created—partially due to the advantages granted to the central bank by the state, but also because the financial system naturally pyramided reserves. As Goodhart (1989b) emphasizes, this development was not intentional but was an unforeseen and evolutionary development. The central bank only very slowly came to recognize that it could exercise some control over the financial system because its liabilities had become the primary reserve as a side-effect of various state policies that were designed to enhance state purchasing power by supporting the development of fiat money. This transition to the modern financial system, which is based on credit money and deposit banking, and in which state purchasing power rests

on fiat money, was essential for the development of the modern capitalist state and economic system.

These developments were largely in place by the 19th century, although the seeds of capitalism were planted earlier. According to Ingham,

The beginning of what Keynes referred to as a “monetary production economy” is to be found in the seventeenth century when signifiers of private debt gradually evolved into widely accepted and legally enforceable means of payment. At this time in western Europe, private bank-issued money existed alongside the sovereign public currencies. ... Eventually the integration of state borrowing and bank lending in the creation of “national” debts led to the creation of entirely new forms of means of payment. (Ingham 2004a, p. 187)

The modern monetary system includes a variety of privately issued debts denominated in the national money of account. These private liabilities form the lower ranks of a pyramid of monetary liabilities, with the sovereign government’s liabilities above them—including treasury bonds and then bills, and central bank notes and then reserves at the top. As Ingham puts it, “Monetary space is a sovereign space in which economic transactions (debts and prices) are denominated in a money of account ... the degree of moneyness is determined by the position of the claim or credit in the hierarchy of acceptability.” (Ingham 2004a, p. 214)

In later chapters, we will examine these relationships in further detail.

NOTE

1. I thank Chris Desan, David Fox, and other participants of a seminar at Cambridge for the discussion I draw upon here.

4. The state theory of money

INTRODUCTION

In the previous chapters, we examined alternative approaches to the nature and origins of money. We saw that historical and anthropological evidence points to an origin in accounting for debts and credits, while the creation of the money of account may have been tied to the rise of authorities. In Chapter 3, we examined the history of private credit money and the development of banking and central banking. We saw that in England this was closely linked to sovereign finance. In this chapter, we examine the state theory of money, emphasizing the state's role in choosing and enforcing a national money of account. We will also explain why the state's money is a form of credit money.

In conventional analysis, money is used to facilitate exchange; its value was originally linked to the value of the precious metal it represented, although under a fiat money system its value is said to be determined by the quantity of commodities it can purchase. This, in turn, is a function of the rate of inflation, presumed to be under the central bank's control. In this view, monetary policy is primarily focused on control of the money supply, while fiscal policy has to do with government spending, taxing, and borrowing.

This is quite different from the Chartalist approach, which can be traced from Adam Smith through to John Maynard Keynes. In this chapter, we first focus on contributions made by Adam Smith, G.F. Knapp, A. Mitchell Innes, and J.M. Keynes, and the related work by Hyman Minsky, Abba Lerner, Kenneth Boulding, and Charles Goodhart. This is the view that informs the analysis of MMT—the State Money or Chartalist approach.

In the Chartalist approach, money is a “creature of the state,” as Lerner put it; at least in the case of the modern period, examples of stateless money are hard to find. The state defines money as that which it accepts at public pay offices (today mainly in payment of taxes). This has important policy implications. Once the state imposes a tax on its citizens, payable in money over which it has a monopoly of issue, it can influence the value of that money by setting the conditions under which the population can obtain it. The government does not need the public's money to spend; rather, the public needs the government's money to pay taxes. This means that the government can buy whatever is for sale in terms of its money by providing that money.

As we discuss in Chapter 8, because the public will normally wish to hold some extra money, the government will usually have to spend more than it taxes; in other words, the normal requirement is for a government deficit. Government deficits do not require borrowing by the government (bond sales); rather (as we show in Chapter 9), the government provides bonds to allow the public to hold interest-bearing alternatives to non-interest-bearing government money. Thus, the Chartalist view of money, if fully understood, would lead to a very different view of appropriate monetary and fiscal policy goals. Most notably, it would be recognized that rather than striving for a balanced budget, deficits would be accepted as the norm. And rather than using monetary policy to pursue stable prices, monetary policy would recognize that its role is to establish the short-term interest rate, while fiscal policy would aim to increase the stability of the value of the currency (see Chapters 8 and 11).

Sometimes the MMT position on State Money has been oversimplified to the point that its conclusions seem to conflict with the historical evidence. For example, if a sovereign government can simply impose a tax payable in its own IOUs, then why are there so many historical examples of governments tying their currency to precious metals, or issuing precious metal coins? If governments can finance all their spending by issuing IOUs, then why do they borrow from private banks and central banks? In the final sections, we will return to the example of the development of the modern monetary system in England to reconcile the theory and the evidence. To put it simply, creating a modern monetary system is easier said than done.

Part of the challenge was to develop an understanding of the nature of money. For thousands of years, there were two main rival views, which can be labeled the metalist and the nominalist views. As discussed earlier, the metalist view is that money derives its value from the commodity used as money (primarily, as a medium of exchange), while the nominalist view is that the material used to record the nominal value of an IOU is a matter of technology. However, it is important to emphasize that the nominalist view was often dominant so it would be a mistake to conclude that the use of precious metal in coins was simply due to a misunderstanding of money's nature. Instead, even sovereign governments that fully understood nominalism were not always able to use that understanding—they were forced to adopt practical metalism.

SMITH ON MONEY

Let us first examine Adam Smith's views of money that are quite similar to the views presented below (especially in Chapter 8)—particularly on bank creation of money and on the determination of the value of an inconvertible currency.

According to Smith, convertible banknotes can substitute for commodity money:

When the people of any particular country have such confidence in the fortune, probity, and prudence of a particular banker, as to believe that he is always ready to pay upon demand such of his promissory notes as are likely to be at any time presented to him; those notes come to have the same currency as gold and silver money ... (Smith [1776] 1937, p. 277)

At this point, the bank can create (bank) money by lending its own notes. In most countries, banknotes enter the economy as banks discount bills of exchange; however, in Scotland, banks had gone one step further:

They invented, therefore, another method of issuing their promissory notes; by granting, what they called, cash accounts, that is by giving credit to the extent of a certain sum ... to any individual who could procure two persons of undoubted credit and good landed estate to become surety for him ... (op. cit., pp. 282–3)

In other words, banks issued notes and held IOUs of borrowers, with the “surety” of two creditworthy persons. These banks would then accept their own notes in payment of bank loans. This ensured demand for banknotes by merchant sellers to make payments on loans (“cash accounts”).

The banks, when their customers apply to them for money, generally advance it to them in their own promissory notes. These the merchants pay away to the manufacturers for goods, the manufacturers to the farmers for materials and provisions, the farmers to their landlords for rent, the landlords repay them to the merchants for the conveniences and luxuries with which they supply them, and the merchants again return them to the banks in order to balance their cash accounts, or to replace what they may have borrowed of them; and thus almost the whole money business of the country is transacted by means of them. (op. cit., p. 283)

Because notes circulate as if they were money, the banker needs to hold only a fractional reserve against them.

Though he has generally in circulation, therefore, notes to the extent of a hundred thousand pounds, twenty thousand pounds in gold and silver may, frequently, be a sufficient provision for answering occasional demands ... [T]he whole circulation may thus be conducted with a fifth part only of the gold and silver which would otherwise have been requisite. (op. cit., p. 277)

Thus, in Scotland, “The business of the country is almost entirely carried on by means of the paper of those different banking companies, with which purchases and payments of all kinds are commonly made.” (op. cit., p. 281) Banknotes “free up” specie that is not needed domestically to go abroad

in order to seek that profitable employment which it cannot find at home. But the paper cannot go abroad; because at a distance from the banks which issue it, and from the country in which payment of it can be exacted by law, it will not be received in common payments. Gold and silver, therefore ... will be sent abroad, and the channel of home circulation will remain filled with ... paper ... (op. cit., p. 278)

Not only does paper money substitute for gold and silver at home, but it also increases the volume of trade. “By means of those cash accounts every merchant can, without imprudence, carry on a greater trade than he otherwise could do.’ (op. cit., p. 283) This is because the merchant with a “cash account” can safely keep nearly zero precautionary balances:

The merchant in Edinburgh ... keeps no money unemployed for answering such occasional demands. When they actually come upon him, he satisfies them from his cash account with the bank, and gradually replaces the sum borrowed with money or paper which comes in from the occasional sales of his goods. (op. cit., p. 284)

This does not mean that the volume of paper money will exceed the volume of gold and silver that would be necessary to circulate the same output: “Should the circulating paper at any time exceed that sum, as the excess could neither be sent abroad nor be employed in the circulation of the country, it must immediately return upon the banks to be exchanged for gold and silver.” (op. cit., p. 284)

Occasionally, banks do issue too much paper money. This could occur because a bank did not require its loans to be repaid; for example, a bank might allow a customer to deliver a bill of exchange rather than either coins or banknotes. These were often “fictitious” bills with no commodities circulating behind them.

[T]he value which had been really advanced upon the first bill, was never really returned to the banks which advanced it; because, before each bill became due, another bill was always drawn to somewhat a greater amount ... than the bill which was soon to be paid; and the discounting of this other bill was essentially necessary towards the payment of that which was soon to be due. This payment, therefore, was altogether fictitious. (op. cit., pp. 295–6)

This process would increase interest owed (due to compounded discounts on the submitted bills) beyond the ability to pay. Further, excessive note issuance would increase redemption, draining reserves and forcing the bank to increase its reserve holdings—which earn less interest, thereby lowering its profitability. Thus, for the most part, market pressures would ensure that there would be a tendency to issue the “correct” amount of paper, equivalent to the quantity of gold and silver required for circulation, but more than the amount that would

have been circulated if specie were actually used in circulation (because the volume of trade would be larger).

So long as paper money is redeemed on demand for gold (or silver), it circulates at par with the gold coin. “Whatever is either bought or sold for such paper, must necessarily be bought or sold as cheap as it could have been for gold and silver.” (op. cit., p. 308) If it is not redeemable on demand, then it might circulate at a discount. He discussed the case where redeemability might be uncertain or might require a wait: “Such a paper money would, no doubt, fall more or less below the value of gold and silver, according as the difficulty or uncertainty of obtaining immediate payment was supposed to be greater or less; or according to the greater or less distance of time at which payment was exigible.” (op. cit., p. 309)

As an example, Smith offered the case of the American colonies, which (he claimed) typically offered conversion only after a wait of several years and did not pay interest on the paper for the waiting period. Still, these colonies passed legal tender laws “to render their paper of equal value with gold and silver, by enacting penalties against all those who made any difference in the price of their goods when they sold them for a colony paper, and when they sold them for gold and silver ...” (op. cit., p. 311) Smith decried such regulations as “tyrannical” and ineffectual, for the colony currency would fall relative to the English pound. However, he also noted that Pennsylvania “was always more moderate in its emissions of paper money than any other of our colonies. Its paper currency accordingly is said never to have sunk below the value of the gold and silver which was current in the colony before the first emission of paper money.” (op. cit., p. 311). Here there is some ambiguity, for he had not previously argued that the depreciation of a non-convertible currency was a function of the quantity of the currency issued, but now he seemed to argue that the more moderate emission of Pennsylvania prevented depreciation.

In the following paragraph, he seems to solve the puzzle. If paper money, whose redeemability is uncertain (or is subject to conditions—such as a waiting period), is accepted in payment of taxes, and if it is not excessively issued relative to the tax liability, then it need not lose value.

The paper of each colony being received in the payment of the provincial taxes, for the full value for which it had been issued, it necessarily derived from this use some additional value, over and above what it would have had, from the real or supposed distance of the term of its final discharge and redemption. This additional value was greater or less, according as the quantity of paper issued was more or less above what could be employed in the payment of the taxes of the particular colony which issued it. (op. cit., p. 312)

Thus, any depreciation would occur because the note issue was well above what was required in payment of taxes. A wiser government could not only

prevent depreciation, it might even cause paper money to carry a premium over specie!

A prince, who should enact that a certain proportion of his taxes should be paid in a paper money of a certain kind, might thereby give a certain value to this paper money; even though the term of its final discharge and redemption should depend altogether upon the will of the prince. If the bank which issued this paper was careful to keep the quantity of it always somewhat below what could easily be employed in this manner, the demand for it might be such as to make it even bear a premium, or sell for somewhat more in the market than the quantity of gold or silver currency for which it was issued. (op. cit., p. 312)

In summary, an essentially non-redeemable paper money could circulate above par on a gold standard if it was legally required by the state in payment of taxes, and if the quantity issued were kept “somewhat below what could easily be employed in this manner.” The key, then, is not really redeemability (for specie—what is also called convertibility), nor is it legal tender laws that attempt to “render their paper of equal value with gold and silver”; rather, it is the acceptance of the paper money in payment of taxes and the restriction of the issue in relation to the total tax liability that gives value to the paper money.

Importantly, Smith recognized that this paper money need not be government fiat currency, for his argument was predicated on the recognition that paper money is the liability of the banking system. All that mattered was that the state accept these banknotes in payment of taxes, in which case they could circulate at par, or even at a premium, relative to specie. Note also that this is the real reason that “paper” remains at home while “specie” can go abroad. If there are gold or silver standards abroad, then specie will always be accepted outside the country since it can be *monetized* or generally accepted in payment of taxes in the foreign country. On the other hand, paper money is denominated only in the domestic unit of account and cannot be monetized or accepted for tax payments abroad.

Finally, while Smith did not explicitly recognize it, the payment of taxes is another form of redemption that removes paper money (and specie) from circulation just as bank money (notes or deposits) is redeemed when notes and cheques are presented for payment or clearing. It is not really convertibility, but rather redemption that removes paper money.

Before concluding this section, we note that recent research by Farley Grubb (2015) has confirmed that the Colonial governments fully understood the points made here. He documents that all the colonies issued paper money that was used for tax payments. The parliaments of each would pass a spending bill that authorized printing a sum of paper notes (say, 10,000 Virginia pounds). At the same time, they would pass a law imposing a Redemption Tax (the explicit use of the word “redemption” indicates the purpose of the tax was to

redeem the currency in tax payments) that was expected to raise the same sum of pounds over a relatively short period of time. From the careful records they kept, Grubb calculates that about three-quarters of the tax would be paid in notes, with the rest paid in British coin (the colonies were not allowed to mint coin—which is why they resorted to printing notes). Once received, the notes were burned (ensuring they could not reenter circulation). The remaining notes continued to circulate—Grubb notes that some preferred to use notes over coins as a medium of exchange.

The general conclusion we can make is that Smith recognized that paper money can retain its nominal value so long as it is *redeemed* by its issuer in payment. Payments of private bank notes to the issuing banks redeem the notes; colonial paper notes are redeemed in tax payments. Redemption ensures that paper money does not need to be backed by precious metal to retain its value. Furthermore, the use of paper money domestically frees the precious metal for other uses, primarily for use in trade. However, as we saw in Chapter 3, trade was largely financed by bills of exchange—not by flows of specie. An important exception was the payment of mercenaries in war and provisioning troops abroad where paper money would not be accepted.

In the next section, we examine Knapp's more general theory of money, which is consistent with, but expands significantly upon, the observations of Smith.

KNAPP: THE STATE THEORY OF MONEY

Georg Friedrich Knapp developed the state theory of money, an approach that is directly opposed to the metalist view, in which the value of money derives from the metal standard (for example, gold or silver) adopted. More generally, according to Knapp, metalists try to “deduce” the monetary system “without the idea of a State.” This, he believes, is “absurd,” for “the money of a state” is that which is “accepted at the public pay offices.” (Knapp [1924] 1973, pp. vii–viii) Knapp's exposition is quite complex and required the creation of a classificatory scheme with dozens of idiosyncratic terms. We will keep the summary simple—to some extent paraphrasing rather than using extensive quotes, for otherwise we would have to define the numerous terms he coined.

According to Knapp, debts are expressed in a unit of value—“the unit in which the amount of the payment is expressed” (op. cit., p. 8) and discharged with means of payment; “a movable thing which has the legal property of being the bearer of units of value.” (op. cit., p. 7) What, then, determines which movable things will act as means of payment to discharge debts? Knapp recognized that means of payment are occasionally changed; sometimes one type of material (say, weighed or coined gold) has been accepted but “suddenly” another (say, weighed or coined silver) takes its place. Therefore, while the

means of payment may be a definite material, it is not bound to any particular material, for it may be changed. (op. cit., pp. 8–25) “A proclamation is made that a piece of such and such a description shall be valid as so many units of value ... Validity by proclamation is not bound to any material. It can occur with the most precious or the basest metals ...” (op. cit., p. 30) He recognized that these transitions require that the state announce a conversion rate (say, so many ounces of gold for so many ounces of silver). Hence, the debts were always nominal and were never actually “metallic”: all debts are converted to the new reference metal, which proves that all units of account must be nominal.

Knapp described the modern monetary system, where Chartal money has developed as a *token*:

When we give up our coats in the cloak-room of a theatre, we receive a tin disc of a given size bearing a sign, perhaps a number. There is nothing more on it, but this ticket or mark has legal significance; it is a proof that I am entitled to demand the return of my coat. When we send letters, we affix a stamp or a ticket which proves that we have by payment of postage obtained the right to get the letter carried. The “ticket” is then a good expression... for a movable, shaped object bearing signs, to which legal ordinance gives a use independent of its material. Our means of payment, then, whether coins or warrants, possess the above-named qualities: they are pay-tokens, or tickets used as means of payment... Perhaps the Latin word “Charta” can bear the sense of ticket or token, and we can form a new but intelligible adjective—“Chartal.” Our means of payment have this token, or Chartal, form. Among civilized peoples in our day, payments can only be made with pay-tickets or Chartal pieces. (Knapp [1924] 1973, pp. 31–2)

Note that, like the tin disc issued by the cloakroom, the material used to manufacture the Chartal pieces is irrelevant—it can be gold, silver, or common metal; it can be paper, and today it can be electronic entries on a hard drive.

It is, therefore, impossible to tell from the pieces themselves whether they are Chartal or not. This is at once evident in the case of warrants. As to coins, we must always refer to the Acts and Statutes, which alone can give information ... if the pieces gain their validity through proclamation, they are Chartal. (op. cit., pp. 34–5)

Finally, “Money always signifies a Chartal means of payment. Every means of payment we call money. The definition of money is therefore a Chartal means of payment.” (op. cit., pp. 34–8)

Chartalism is often identified with the proposition that legal tender laws determine what must be accepted as means of payment (following Schumpeter’s interpretation of Knapp). However, Knapp’s analysis went further:

If we have already declared in the beginning that money is a creation of law, this is not to be interpreted in the narrower sense that it is a creation of jurisprudence,

but in the larger sense that it is a creation of the legislative activity of the State, a creation of legislative policy. (op. cit., p. 40)

And what is the nature of this “legislative activity” that determines what will be the Chartalist money accepted within the jurisdiction of the state?

What forms part of the monetary system of the State and what does not? We must not make our definition too narrow. The criterion cannot be that the money is issued by the State, for that would exclude kinds of money which are of the highest importance; I refer to bank-notes: they are not issued by the State, but they form a part of its monetary system. Nor can legal tender be taken as the test, for in monetary systems there are very frequently kinds of money which are not legal tender ... We keep most closely to the facts if we take as our test, that the money is accepted in payments made to the State's offices. *Then all means by which a payment can be made to the State form part of the monetary system.* On this basis it is not the issue, but the acceptance, as we call it, which is decisive. State acceptance delimits the monetary system. By the expression “State-acceptation” is to be understood only the acceptance at State pay offices where the State is the recipient. (op. cit., p. 95; emphasis added)

It is the decision of the state to accept at state pay offices, and not legal tender laws, that creates a Chartal money. As discussed earlier in Chapter 3, if the state joins the bank *giro*, its currency will circulate along with bank money; if the government accepts the bank *giro* money, then bank money is part of the state's monetary system.

But among those Chartal moneys that are part of the monetary system, there must be a special one:

In the monetary system of a State there must be one kind of money which is definitive, as opposed to provisional (convertible) money ... Money is definitive if, when payment is made in it, the business is completely concluded ... The payer is no longer under an obligation, the recipient has no further rights either against the payer or against the State, if the State has issued the money. (op. cit., p. 102)

That kind of definitive money which is always kept ready and can be insisted on for *apocentric payments* [payments made by the state]...we call *valuta*, all other kinds of money ... we call accessory. (p. 105)

Thus, it is the decision of the state to accept at state pay offices, and not legal tender laws, that creates a Chartal money.

According to Knapp, *centric* payments, or those involving the state, are decisive; these take the form of either (1) “payments to the State as receiver; these we call *epicentric*” or (2) “payments made by the State, these we will call apocentric.” (ibid., pp. 96–7) On the other hand, payments between private persons (*paracentric*) “are not so important as is generally supposed, for they

mostly, so to speak, regulate themselves.” (op. cit., p. 96) Indeed, the actions of the state play a large role in determining what serves as *paracentric* means of payment in private transactions. The *definitive* money is that which the state will accept at pay offices, while *valuta* money is a component of definitive money, namely what it provides in payment:

In Germany our gold pieces were *valuta*, not because they were made of gold ... but only because the State, when it made a payment, was ready in the last resort to pay in gold pieces, and, if it found it at all inconvenient, totally to refuse any other means of payment which the recipient might happen to want. (op. cit., p. 107)

However, once the state has decided to declare one type of money as *valuta*, that type will become the decisive money used in private transactions:

So, if from political necessity the State announces that henceforth it will pay in State notes, as fountain of law it must equally allow the State notes to suffice for other payments ... The consequence is, in a legal dispute the means of payment which the creditor is compelled to accept is always that which the State has put in the position of *valuta* ... Apart from friendly agreement, all payments eventually have to be made in *valuta* money. (op. cit., p. 110)

It is not simply a legal tender law that makes state notes acceptable in private transactions, but the fact that the state first decides what it will use or accept as money in its own transactions, and that this must then be acceptable as a means of settlement of private debts. “The laws do not decide what shall be *valuta* money, they merely express a pious hope, for they are powerless against their creator, the State ...” (op. cit., p. 111)

Knapp extended his analysis to include bank money:

The bank makes notes and offers them in payment to its customers. Issuing notes is not a special business... but a special way in which the bank endeavours to make its payments ... It tries to pay in its own notes instead of in money issued by the State, because then with a comparatively small capital it can make greater profits than it otherwise could. (op. cit., p. 131)

Acceptability of banknotes in private transactions is not (as was commonly believed) due to the bank’s promise to convert these to specie: “An inconvertible bank-note, then, is not a nullity, but has this in common with the convertible banknote, that it is a till-warrant of the bank.” (op. cit., p. 134) What is important is that the note “is a private till-warrant available for payments to the bank ... but clearly the customers of the bank can use it for payments between themselves, as they are sure it will be taken at the bank. These customers and the bank form, so to speak, a private pay community; the public pay community is the State.” (op. cit.)

Within the “private pay community” (or *giro*), bank money is the primary money used in payments; however, payments in the “public pay community” require state money. This can include bank money, but note that generally, the delivery of bank money to the state is not final, or definitive, because the state will present it to banks for “redemption” (for *valuta* reserves). Bank money, when used in the public pay community, is not definitive unless the state also uses it in its own purchases.

What can make banknotes state money? Banknotes are not automatically money of the state, but they become so as soon as the state announces that it will receive them in *epicentric* payments (payments to the state). (op. cit., p. 135) If the state accepts notes in payment to the state, then the banknotes become *accessory* and the business of the bank is enhanced, “for now everybody is glad to take its banknotes since all inhabitants of the state have occasion to make *epicentric* payments (e.g., for taxes).” (op. cit., p. 137) The banknotes then become *valuta* money if the state takes the next step and makes “*apocentric payments* [payments by the state] in banknotes.” (op. cit., p. 138)

However, states often required that banks make their notes convertible to state-issued money: “one of the measures by means of which the State assures a superior position to the money which it issues itself” (op. cit., p. 140), and thus maintained banknotes in the role of *accessory* money (rather than allowing them to become *valuta* money). If the state accepts banknotes in payment but does not make payments in these banknotes, then the notes will be redeemed—leading to a drain of reserves (indeed, governments and central banks use redemption or the threat of redemption to discipline banks).

In times of distress, however, governments would pass laws ending convertibility, announce that the state would henceforth make payments in terms of the banknotes, and declare that the banknotes were *valuta* money. (op. cit., p. 143) Sometimes this was for one bank only—the bank which became the central bank. Through the action of the state, then, paper money can become *valuta* money. “At first bank-notes and Treasury notes are employed only as *accessory* money ... The mournful hour arrives when the State has to announce that it can no longer pay in the money that was till then *valuta* [say, coined gold] and that those warrants themselves are now *valuta*.” (op. cit., p. 196)

At this point, we have a Chartalist, non-convertible, paper money—often called “fiat money”—as in the case of modern developed countries. This extreme development came nearly three-quarters of a century after Knapp’s book was first published (1905)—when we finally abandoned Bretton Woods. However, he had recognized that the money of a state does not derive its value from metal, and indeed, that no metal is needed domestically. On the other hand, in the international sphere, “To dispense with specie money altogether would only be possible for very large federations of States [and, therefore, is] probably impracticable. On account of foreign trade specie money is still

necessary.” (op. cit., p. xv) Within a state, however, specie is not necessary, for “state money may be recognized by the fact that it is accepted in payment by the State”; as Keynes said (see below), the state not only enforces the dictionary (legal tender laws) but writes it (decides what is to be accepted as money).

We can see the development of the gold standard as a solution to the problem of what to use for international—stateless—money. The problem is that tying a domestic currency to gold created issues of internal monetary stability even as it resolved the need for international money. As it happened, however, neither the gold standard nor the Bretton Woods standard was sustainable because internal instability generates external instability.

Once freed from domestic convertibility on a metallic standard, the state’s spending domestically would not be constrained by the quantity of the metal available. Abandonment of the metallic standard internationally would eliminate metallic constraints on countries. The state thus moved to a paper money system domestically, making its *apocentric* payments in central bank notes and accepting *epicentric* payments in private bank notes (today, deposits) that would have to be redeemed (today, cleared) for the *valuta* central bank notes (today, reserves). Precious metals were then used only for international purposes until the US finally abandoned the gold standard altogether in the early 1970s. Today, it is primarily the US Dollar that is used as the international reserve (*valuta*) money.

Discussion of the international monetary system necessarily takes us beyond state money because sovereign governments cannot generally use their sovereign power to impose liabilities in foreign nations. We next turn to Innes, who integrated state and credit monies.

INNES: AN INTEGRATION OF THE STATE AND CREDIT APPROACHES TO MONEY

A. Mitchell Innes (1913, 1914) suggested that we can locate the origins of credit and debt relations in the elaborate system of tribal *wergild* designed to prevent blood feuds. (See also Ingham 2004a, 2004b, 2013a, 2013b; Grierson 1977, 1979; and Goodhart 1998.) As discussed in Chapter 2, *wergild* fines were paid by transgressors directly to victims and their families and were established and levied by public assemblies. A list of fines for each transgression was developed, and a designated “rememberer” would be responsible for passing it down to the next generation. Note that each fine was levied in terms of a particular good that was both useful to the victim and more-or-less easily obtained by the perpetrator.

Innes believed that money evolves not from a pre-money market system but rather from the *penal system*, and he highlights the important role played by authorities in the origins and evolution of money. (See also Ingham 2004b and

Grierson 1977 for similar arguments.) More specifically, the state (or any other authority able to impose obligations) imposes a liability in the form of a generalized, social unit of account—money—used for measuring the obligations. This does not require the pre-existence of markets and, indeed, almost certainly predates them, resolving the conundrum faced by methodological individualists and emphasizing the social nature of money and markets—which did not spring from the minds of individual utility maximizers, but rather were socially created.

The state chooses the unit, names what is accepted in payment of obligations to itself, and (eventually) issues the token or “money-thing” (in Keynes’s terminology—see below) it accepts. The material from which the money issued by the state is produced is not important (whether it is coined gold or base metal, paper, or, now, digitized numbers at the central bank). No matter what it is made of, the state must announce its nominal value (at a minimum, the value at which the money is accepted in meeting obligations to the state) and accept it in payments made to the state.

What is most interesting about the contributions of Innes is that he integrated the state theory of money with a credit theory of money. Along the lines of the latter, Schumpeter distinguished between the “monetary theory of credit” and the “credit theory of money.” The first sees private “credit money” as only a temporary substitute for “real money”—possibly a “natural” money that is free of social relations, i.e., a commodity. Final settlement must take place in real money, which is the ultimate unit of account, store of value, and means of payment. Exchanges might take place based on credit, but credit expansion is strictly constrained by the quantity of real money. Ultimately, only the quantity of real money matters as far as economic activity is concerned. Until recently, modern macroeconomic theory was based on the concept of a deposit multiplier (see Chapter 5) that links the quantity of privately created money (mostly, bank deposits) to the quantity of high-powered money (HPM, also called monetary base—central bank reserves plus currency). This is the modern equivalent to what Schumpeter called the monetary theory of credit, and Friedman (or Brunner) is the best representative. The real money that is the basis of deposit expansion should be controlled, preferably by a rule that will make the modern fiat money operate more like the metallic money of the hypothesized past.

Schumpeter’s credit theory of money, by contrast, emphasizes that credit expands to allow economic activity to grow. Credit creates new claims on HPM even as it leads to new production. However, because there is a clearing system that cancels claims and debits without the use of HPM, credit is not merely a temporary substitute for HPM. Schumpeter does not deny the role played by HPM as an ultimate means of settlement; he simply denies that it is required for most final settlements.

Like Schumpeter, Innes focused on credit and the clearing system, mocking the view that “in modern days a money-saving device has been introduced called credit and that, before this device was known all purchases were paid for in cash, in other words in coins.” (Innes 1913, p. 389) Instead, he argued, “careful investigation shows that the precise reverse is true.” (Innes 1913, p. 389) Rather than selling in exchange for “some intermediate commodity called the ‘medium of exchange,’” a sale is really “the exchange of a commodity for a credit.” Innes called this the “primitive law of commerce”: “The constant creation of credits and debts, and their extinction by being canceled against one another, forms the whole mechanism of commerce...” (Innes 1913, p. 393) Innes explains:

By buying we become debtors and by selling we become creditors, and being all both buyers and sellers we are all debtors and creditors. As debtor we can compel our creditor to cancel our obligation to him by handing to him his own acknowledgment [sic] of a debt to an equivalent amount which he, in his turn, has incurred. (Innes 1913, p. 393)

The market, then, is not viewed as the place where goods are exchanged, but rather as a clearing house for debts and credits. Indeed, Innes rejected the typical analysis of the medieval village fairs, arguing that these were first developed to settle debts, with retail trade later developing as a sideline to the clearing house trade. On this view, debts and credits and clearing are the general phenomena; trade in goods and services is subsidiary—one of the ways in which one becomes a debtor or creditor, or clears debts. Innes viewed the creditor-debtor relation as the fundamental social relation behind money’s veil. There is no natural relation-free money that lies behind the credit money. Indeed, for Innes even HPM is credit money, linking the sovereign and the public in a debt-credit relationship.

The credit approach as advanced by Innes and Schumpeter provides a more useful vision of the monetary operations of a capitalist, “market” economy than does the orthodox vision of money serving as a lubricating medium of exchange. The monetary production economy, as described by Marx, Veblen, and Keynes, is dominated by a complex web of financial relations that were characterized by Minsky as “money now for money later” propositions. (See below, as well as Minsky 1986 p. 228) Money is not a veil that should be stripped away to observe the essential characteristics of the “market economy.” Rather, the money of account and those credit-debt relations are the key institutional relations of the capitalist economy.

Innes sounds much like Knapp on the state’s money, although there is no direct indication that he was drawing on Knapp’s exposition. However, Innes is best at generalizing the credit approach to money to include the state’s own

currency as a credit money. He insisted that when the state spends, it becomes a debtor (as he said, “by buying we become debtors”) as it issues state money. Hence, even state money is credit money; however, it is a special kind of credit, “redeemed by taxation.” (1914, p. 168) For the government, a dollar is a “promise to pay,” a promise to “satisfy,” a promise to “redeem,” just as all other money is. (Our word, pay, derives from “pacify”—with an obvious link to the meaning of “satisfy.”) Innes argued that even on a gold standard it is not gold that the government promises to pay, for while it is true that the government paper money is in that case convertible into gold coin, “redemption of paper issues in gold coin is not redemption at all, but merely the exchange of one form of obligation for another of an identical nature.” (Innes 1914, p. 165)

Whether the government’s IOU is printed on paper or on a gold coin, it is indebted. What, then, is the nature of the government’s IOU? Innes identifies the “very nature of credit throughout the world,” which is “the right of the holder of the credit (the creditor) to hand back to the issuer of the debt (the debtor) the latter’s acknowledgment or obligation.” (1914, p. 161) The holder of a coin or certificate has the right to pay any debt due to the government by tendering that coin or certificate, and it is this right and nothing else which gives it value. (ibid.) Hence, we can integrate the state money and credit money approaches through the recognition of the “very nature of credit,” which is that the issuer must accept its own IOUs.

What, then, is special about government? The government’s credit

...usually ranks in any given city slightly higher than does the money of a banker outside the city, not at all because it represents gold, but merely because the financial operations of the government are so extensive that government money is required everywhere for the discharge of taxes or other obligations to the government. (Innes 1914, p. 154)

The special characteristic of government money, then, is that it is “redeemable by the mechanism of taxation” (Innes 1914, p. 15): “[I]t is the tax which imparts to the obligation its ‘value’ ... A dollar of money is a dollar, not because of the material of which it is made, but because of the dollar of tax which is imposed to redeem it.” (Innes 1914, p. 152)

In spite of the attention paid to the gold standard, it was only intermittently in place for relatively short periods—with the exception of England after Queen Elizabeth’s reform discussed in Chapter 3. Typically, the money issued by the authorities has not been gold money, nor was there any promise to convert the sovereign’s money to gold. Indeed, as Innes insisted, throughout most of Europe’s history, the money issued by the state took the form of the hazelwood tally stick: “This is well seen in medieval England, where the regular method used by the government for paying a creditor was by ‘raising a tally’ on the

Customs or on some other revenue-getting department, that is to say, by giving to the creditor as an acknowledgment of indebtedness a wooden tally.” (Innes 1913, p. 398; see also Maddox 1969) Other money forms included records on clay tablets, leather and base metal coins, and paper certificates.

Why would the population accept otherwise “worthless” sticks, clay, base metal, leather, or paper? Because these were evidence of the State’s liabilities that it would accept in payment of taxes and other debts owed to itself. The key power of the state is its ability to impose taxes:

[t]he government by law obliges certain selected persons to become its debtors ... This procedure is called levying a tax, and the persons thus forced into the position of debtors to the government must in theory seek out the holders of the tallies or other instrument acknowledging a debt due by the government. (Innes 1913, p. 398)

Contrary to orthodox thinking, the desirability of the money issued by the state was not determined by intrinsic value (even on the gold standard), but by the nominal value set by the state at its own pay offices. Nor was the government’s money necessarily forced onto the public through legal tender laws. It is certainly true that governments often adopt legal tender laws, but these are difficult to enforce and hence often ineffective. (Knapp 1924, p. 111) The power of the government to impose a tax and to name what it accepts in tax payment is sufficient, trumping legal tender laws.

Once the state has created the money of account and named what can be delivered to fulfill obligations to the state, it has generated the necessary preconditions for the development of markets. (Ingham 2004a, p. 25) Credits and debts preceded markets, and indeed, created the need for markets. The tax obligation creates the incentive for private credits and debts denominated in the money of account and then for markets with prices denominated in the same money of account. Indeed, evidence from early Babylonia suggests that early authorities set prices for each of the most important products and services—perhaps those accepted to meet obligations to the authorities. Once prices in money were established, it was a short technical leap to the creation of markets. This stands orthodoxy on its head by reversing the order: first money and prices, then markets and money as a medium of exchange (rather than barter-based markets and relative prices, and then *numeraire* money and nominal prices). The next step was the recognition by the government that it could issue currency to purchase the mix it desired, then receive the same currency in the tax payments by subjects/citizens. This would further the development of markets because those with tax liabilities but without the goods and services the government wished to buy would have to produce for market to obtain the means of paying obligations to the state.

Today, it is usually only the sovereign government that can impose liabilities on others (although in the past, feudal lords and conquerors could do so). This puts it in a privileged position because it can create a demand for its own liabilities by requiring that taxpayers deliver government liabilities as payment of taxes. It can also enact legal tender laws and legal reserve requirements to provide further privilege to treasury and central bank liabilities. Finally, the modern state is, of course, a very large entity—hence an important purchaser of output and source of income—which makes its liabilities ubiquitous. Still, if the state did not impose tax liabilities in its currency and require ultimate payments to itself in the form of its treasury and central bank liabilities, it is difficult to believe that its sheer size and its legal tender laws alone would be sufficient to guarantee its current spot at the top of the money hierarchy.

In sum, Innes offered an integration of the state money and credit money approaches. While some of his arguments may have been too strong—and too simplistic—they are largely consistent with those made by Smith and Knapp and generally in line with the discussion of the history of money in Chapter 3. As we saw with the close study of the development of the modern monetary system in England, it is not quite so simple as Innes implies for a sovereign authority to use its tax authority to ensure its fiat money will be accepted. That history was more complicated, with the Crown relying on both a strong metallic currency and support of London's moneyed interests. However, with the rise of modern states, central banks, and nominal democracies, we have achieved systems that more or less operate along the lines suggested by Knapp and Innes.

We now turn to Keynes's views.

KEYNES: THE PRIMACY OF THE MODERN STATE

While Keynes's *General Theory* presented the theory of aggregate effective demand identified as "Keynesian theory," his earlier *Treatise on Money* provided a more detailed treatment of his monetary theory. (Keynes (1937) 1964, Keynes (1930) 1976) He was heavily influenced by both Knapp and Innes and indeed played a role in promoting their work. He reviewed the 1913 article by Innes in the *Economic Journal*, arguing that while further research might call into question some of the claims about the history of money made by Innes, Keynes said his understanding of money appeared sound. Apparently, this led to a phase Keynes called his "Babylonian Madness," during which he studied ancient monies and metrology, hence his emphasis on money as a measure (unit of account)—see Ingham (2004a). This influenced his *Treatise* a decade later. Further, Keynes played a role in getting Knapp's book translated into English in 1924. There is some question about whether Keynes could read

German, but he could certainly read the translation after 1924. In important respects, his exposition in the *Treatise* follows that of Innes and Knapp.

According to Keynes, the *money of account* is the *primary concept* of a theory of money; it “comes into existence along with Debts, which are contracts for deferred payment, and Price-Lists, which are offers of contracts for sale or purchase” (Keynes 1976, p. 3). In turn,

Money itself, namely that by delivery of which debt-contracts and price contracts are discharged, and in the shape of which a store of General Purchasing Power is held, derives its character from its relationship to the Money-of-Account, since the debts and prices must first have been expressed in terms of the latter. (ibid.)

He further clarifies the distinction between money and money of account: “the money-of-account is the description or title and the money is the thing which answers to the description.” (op. cit., pp. 3–4) Following Knapp, Keynes argued that the state determines what serves as the money of account as well as dictates what “thing” will be accepted as money:

The State, therefore, comes in first of all as the authority of law which enforces the payment of the thing which corresponds to the name or description in the contracts. But it comes in doubly when, in addition, it claims the right to determine and declare what thing corresponds to the name, and to vary its declaration from time to time—when, that is to say, it claims the right to re-edit the dictionary. This right is claimed by all modern states and has been so claimed for some four thousand years at least. (Keynes 1976, p. 4)

The “Age of Chartalist or State Money” was reached when the state “claimed the right not only to enforce the dictionary but also to write the dictionary.” (op. cit., p. 5) Let us emphasize that Keynes believed the “Age of State Money” to have begun “at least” four thousand years ago. As such, the state theory of money would apply to all “modern” economies, including those living under the gold standard in the last century—even a gold-based “commodity” money is state money. We do not know whether money was used in prehistoric times, so its nature might have been different, but the age of modern money began with the rise of authorities at least four thousand years ago. As an aside, the title of the *Understanding Modern Money* book (Wray 1998), comes from this statement—MMT applies to the past four thousand years, at least.

Privately issued debt—such as bank money—might be accepted in the settlement of transactions even if it is not declared by the government to be money; it can circulate “side by side” with “state money.” (op. cit., p. 6) However, the state might “use its chartalist prerogative to declare that the [bank] debt itself is an acceptable discharge of a liability.” (ibid.) Bank money then becomes “Representative Money.” (ibid.)

At the cost of not conforming entirely with current usage, I propose to include as State-Money not only money which is itself compulsory legal-tender, but also money which the State or the central bank undertakes to accept in payments to itself or to exchange for compulsory legal-tender money. (ibid.)

In a footnote to this passage, he goes on: “Knapp accepts as ‘Money’ – rightly I think – anything which the State undertakes to accept at its pay-offices, whether or not it is declared legal tender between citizens.” (op. cit., pp. 6–7) Therefore, like Knapp, Keynes’s analysis goes beyond legal tender laws to identify state *acceptation* as the key to determining what will serve as money.

Finally, state money may take any of three forms: “*Commodity Money*, *Fiat Money* and *Managed Money*, the last two being sub-species of *Representative Money*.” (op. cit., p. 7) Commodity money is defined as “actual units of a particular freely-obtainable, non-monopolized commodity which happens to have been chosen for the familiar purposes of money,” or “warehouse warrants for actually existing units of the commodity.” (ibid.) Fiat money is representative money “which is created and issued by the State but is not convertible by law into anything other than itself and has no fixed value in terms of an objective standard.” (ibid.) This is distinguished from managed money, which “is similar to Fiat Money, except that the State undertakes to manage the conditions of its issue in such a way that, by convertibility or otherwise, it shall have a determinant value in terms of an objective standard.” (op. cit., p. 8)

Managed money is, according to Keynes, the most general form of money, which can “degenerate into Commodity Money on the one side when the managing authority holds against it a hundred percent of the objective standard, so that it is in effect a warehouse warrant, and into Fiat Money on the other side when it loses its objective standard.” (ibid.) Note that Keynes argued that even a gold standard operates as a state money system because, in either case, the state can always “rewrite the dictionary,” for example, by adopting a silver standard and a conversion rate (say, one ounce of gold for four ounces of silver). This is the pyramidal system based on the Bank of England examined in Chapter 3.

State money can be held by banks, by the central bank, and by the public:

The State-Money held by the central bank constitutes its “reserve” against its deposits. These deposits we may term Central Bank-Money. It is convenient to assume that all the Central Bank-Money is held by the Member Banks – in so far as it may be held by the public, it may be on the same footing as State-Money or as Member Bank-Money, according to circumstances. This Central Bank-Money plus the state money held by the Member Banks makes up the Reserves of the Member Banks, which they, in turn, hold against their Deposits. These Deposits constitute the Member Bank-Money in the hands of the Public, and make up, together with the State-Money (and Central Bank-Money, if any) held by the Public, the aggregate of Current Money. (Keynes 1976, pp. 9–10)

Any payments to the state using “Member Bank-Money” will cause member banks to lose “Central Bank-Money” or “State Money held by the Member Banks”—that is, reserves.

As Knapp recognized, “Member Bank-Money” is the primary “thing” answering to the “description”—money—used in private transactions (or, within the “private pay community”—the private *giro*). When accepted in payment of taxes, it is also used in the “public pay community” (or public *giro*), but it is not “*definitive* or *valuta* money from the perspective of member banks, because they must deliver reserves (mainly “Central Bank-Money”) whenever taxes are paid using “Member Bank-Money.”

In summary, with the rise of the modern state, the money of account (“the description”) is chosen by the state, which is free to choose what will qualify as money (the “thing” that answers to the description). This goes beyond legal tender laws, which establish what can legally discharge contracts, to include that which the state accepts in payment at its “pay offices.” The state is free to choose a system based on commodity money, fiat money or managed money. Even if it chooses a strict commodity system, the value of the money does not derive from the commodity accepted as money,

[f]or Chartalism begins when the State designates the objective standard which shall correspond to the money-of-account. (Keynes 1976, p. 11)

[M]oney is the measure of value, but to regard it as having value itself is a relic of the view that the value of money is regulated by the value of the substance of which it is made, and is like confusing a theatre ticket with the performance. (Keynes 1983, p. 402)

Once it is recognized that the state “writes the dictionary,” it becomes obvious that the nominal value of a commodity (or managed) money cannot be derived from the value of the “objective standard”; it is then a small step to a “fiat money” with no “objective standard,” for in all these cases, the state determines the nominal value of money by establishing what it will accept at public pay offices, as well as the nominal value of the *thing* accepted.

It is important to understand that Keynes’s approach to money underlies his *General Theory*, which he began to work on almost immediately after the publication of the *Treatise*. He did not revisit these arguments, arguing that the monetary details fall into the background. However, the earliest drafts of the *GT* explicitly proposed to develop a *monetary theory of production*—a theory of the role money plays in the capitalist economy. While the *GT* is largely remembered for its explication of Keynes’s theory of effective demand, the importance of money is downplayed or even reduced to the role it plays in determining “the” interest rate. We will return to Keynes’s views on the role that money plays in the capitalist economy in Chapters 5, 7, and 10.

RECENT CONTRIBUTIONS IN THE CHARTALIST TRADITION

In recent years, many theorists have contributed to the development of approaches consistent with the Chartalist position. Among these, we could include the revival of the endogenous money view that sees money as credit, as well as the Circuit approach that also focuses on the private creation of credit money. Central bank reserves are not a constraint but are used for clearing. It is only in the 20th century that an overwhelming majority of economists came to accept the *exogenous* money view, which posits that the central bank can directly control the quantity of money and that the money stock can be taken to be *fixed* such that it does not respond to “money demand.” That will be taken up in the next chapter. In the remainder of this section, we examine only briefly contributions directly related to arguments made above. However, it is important to note that if money supply responds to money demand, this means that the quantity of money is not “exogenous” in the sense of being determined either through monetary policy (such as control by the central bank over bank reserves) or by the quantity of a precious metal reserve (as under a *commodity money* or *managed money* system). While the state defines money, it does not control the quantity.

The state may be able to control its initial emission of currency, but this is done through fiscal policy rather than monetary policy. That is, the quantity of currency created is partially determined by purchases of the state (including goods, services, and assets purchased by the Treasury and the central bank); much of this currency will then be removed from circulation as taxes are paid. The rest ends up in desired hoards or flows to banks to be accumulated as bank reserves. Monetary policy then drains excess reserves, removing them from member bank accounts and replacing them with bonds voluntarily purchased—as will be discussed in detail in Chapter 9. As Boulding (1950) had argued, fiscal policy has more to do with the quantity of money issued by the government, while monetary policy has to do with the regulation of financial markets (most importantly, with the determination of short-term interest rates—points we return to in Chapters 8 and 9).

Hyman Minsky presented a view of money that was based on the Chartalist approach. As discussed in the next chapter, he emphasized the endogeneity of money, that is, the view that money is created during the normal, and important, processes of a capitalist economy—and is not created and dropped by helicopters (as in Friedman’s famous exogenous, helicopter money story). For the most part, bank money is created as banks make loans:

Money is unique in that it is created in the act of financing by a bank and is destroyed as the commitments on debt instruments owned by banks are fulfilled. Because

money is created and destroyed in the normal course of business, the amount outstanding is responsive to the demand for financing. (Minsky, 1986, p. 249)

A loan is nothing more than an agreement by a bank to make payments now on the basis of a promise by the borrower to pay later: “Loans represent payments the bank made for business, households and governments in exchange for their promises to make payments to the bank at some future date.” (ibid., p. 230)

All of this occurs on the balance sheets of banks; the money that is created by a bank is nothing but a credit to another bank’s balance sheet. According to Minsky, there is a pyramid of liabilities, with those of the central bank at the top. Bank liabilities are convertible on demand into central bank liabilities, which are used for interbank clearing (as well as for the conversion of bank liabilities to “cash” held by the public, resulting in a net reserve drain).

The payments banks make are to other banks, although they simultaneously charge the account of the customer. In the receiving bank, the payments are credited to a depositor’s account ... For member banks of the Federal Reserve System, the interbank payments lead to deposits shifting from the account of one bank to the account of another at Federal Reserve banks. For nonmember banks, another bank – called a correspondent – intervenes, so that the transfer[s] at the Federal Reserve banks are for the accounts of correspondents. (Minsky 1986, pp. 230–1)

Thus, payments among banks occur on the balance sheet of the Fed as banks use central bank money (reserves) to settle net debits from their accounts. “Whereas the public uses bank deposits as money, banks use Federal Reserve deposits as money. This is the fundamental hierarchical property of our money and banking system.” (op. cit., p. 231) This is the same hierarchical arrangement noted by Knapp (in his public and private pay communities) and by Keynes above, and is consistent with the arrangement in England after the development of the banking system under the Bank of England, as discussed in Chapter 3.

In an argument very similar to Knapp’s Chartalist view, Minsky explained that people accept bank money in part because they can use it to meet their own commitments to banks. “Demand deposits have exchange value because a multitude of debtors to banks have outstanding debts that call for the payment of demand deposits to banks. These debtors will work and sell goods or financial instruments to get demand deposits.” (Minsky 1986, p. 231) In other words, according to Minsky, bank money has (nominal) value precisely because it can be used to retire debts to banks—it is, so to speak, accepted at “bank pay offices.” The borrower retires his/her promise to the bank by delivering bank liabilities at a future date, and the need for bank liabilities to retire one’s own liabilities to banks leads one to accept bank liabilities in payment for goods and services delivered. This is also consistent with the universal law

of debt of Innes: an entity must accept its own liability in payment. Rather than focusing on money as a medium of exchange, this focus is on money as a means of payment—to retire liabilities through redemption.

This led Minsky back to the Smith/Knapp/Innes recognition that taxes give value to the money issued by the government.

In an economy where government debt is a major asset on the books of the deposit-issuing banks, the fact that taxes need to be paid gives value to the money of the economy ... [T]he need to pay taxes means that people work and produce in order to get that in which taxes can be paid. (Minsky 1986, p. 231, fn 10)

And even though most taxes are actually paid using bank money, because of the hierarchical arrangement, Keynes and Minsky emphasize that banks can make these payments to the government only by using central bank money, that is, by losing reserves.

Returning to the primary Chartalist theme, Abba Lerner insisted

[W]hatever may have been the history of gold, at the present time, in a normally well-working economy, money is a creature of the state. Its general acceptability, which is its all-important attribute, stands or falls by its acceptability by the state. (Lerner 1947, p. 313)

Just how does the state demonstrate acceptability?

The modern state can make anything it chooses generally acceptable as money ... It is true that a simple declaration that such and such is money will not do, even if backed by the most convincing constitutional evidence of the state's absolute sovereignty. But if the state is willing to accept the proposed money in payment of taxes and other obligations to itself the trick is done. Everyone who has obligations to the state will be willing to accept the pieces of paper with which he can settle the obligations, and all other people will be willing to accept these pieces of paper because they know that the taxpayers, etc., will accept them in turn. (ibid.)

This is as clear a statement of Chartalism as one can find: even if it has not always been the case, it surely is now obvious that the state establishes the “description” of money when it denominates the tax liability in a money of account and defines the “thing” that “answers to the description” when it decides what will be accepted at public pay offices. The “thing” answering the “description” is widely accepted not because of sovereignty alone, not because of legal tender laws, and not because it might have (or have had) gold backing, but because the state has the power to impose and enforce tax liabilities and the right to choose “that which is necessary to pay taxes.” This right, as emphasized by Keynes, “has been so claimed for some four thousand years at least.” (Keynes 1976, p. 3) While one might quibble over the exact number of

years since states first claimed these rights, and while it may have been difficult to establish these rights, there can be no doubt that all modern states do have them.

States can—and sometimes do—lose them. In a sense, the nations that joined the Euro voluntarily gave up their right to choose their currency. Some countries have been forced to abandon their currency as a result of political and economic catastrophe. As Lerner said, “Cigarette money and foreign money can come into wide use only when the normal money and the economy in general is in a state of chaos.” (Lerner, 1947, p. 313) One might only add that when the state is in crisis and loses legitimacy, and in particular loses its power to impose and enforce tax liabilities, *normal money* will be in a *state of chaos*, leading, for example, to the use of foreign currencies in private domestic transactions. In all other cases, it is state money that is used, and state money is what the state accepts in payment of taxes.

CONCLUSION: THE STATE THEORY OF MONEY VERSUS THE METALIST APPROACH

Goodhart (1998) makes a useful distinction between the *metalist* approach and the *chartalist*—or state money—approach. (See also Lau and Smithin 2002, who prefer the term *catallactic* over metalist.) The state money approach highlights the important role played by authorities in the origin and evolution of money. More specifically, the state (or any other authority able to impose an obligation) imposes a liability (fees, fines, taxes, tribute) in the form of a generalized, social unit of account—a money—used for measuring the obligation. Once authorities can levy such obligations, they can name what can be delivered to satisfy them. Initially, the payment might be made “in-kind” by delivering a goat or a bushel of barley, with the authority setting the monetary value of each. This establishes a money price for the goat and the barley, which also sets the relative price between the two. As Hudson argues, price lists were posted to temple columns and those prices would also come to be used in private exchanges. Rather than prices springing naturally out of markets, they were socially created.

The state money approach might appear to be inconsistent with the credit money approach described in the previous chapter. Indeed, some critics of the state money approach—including the metalists—imagine that markets operating on the basis of private credits and debits denominated in a money of account pre-existed the state or authorities. The state is then supposed to have inserted itself into the private money system, taxing and borrowing the private credit money for use in public expenditures. In contrast, Innes insisted that when the state spends, it becomes a debtor (as he said, “by buying we become debtors”) as it issues state money redeemable in payment of taxes. (Innes 1914,

p. 168) State money is a credit money, too. The difference is that state money is redeemed in taxation, while bank money is redeemed in repayment of bank loans, or by conversion of deposits to state money (either cash or reserves).

In the Chartalist approach, the public demands the government's money because that is the form in which taxes are paid ("taxes drive money"). It is not a coincidence that the modern state uses the same *valuta* money in its *apocentric* payments that it accepts in *epicentric* payments—it uses taxes to induce the population to supply goods and services to the state, supplying in return the money that will be used to retire the tax liability. In the modern economy, it appears that taxes are paid using bank money, but analysis of reserve accounting shows (as demonstrated in Chapter 9) that tax payments always lead to a reserve drain (that is, reduce central bank liabilities), so that only the government's money is definitive (finally discharging the tax liability). However, in modern economies the public and private *giros* are intertwined, following the example of England, in mutually reinforcing ways.

We return to the question often raised as an objection to Chartalism: if it is all so simple, why did it take so long for states to move to fiat currency? Why did they coin gold or peg their currency to gold? Why did they borrow from banks? Why couldn't they simply impose a tax and then issue a fiat currency to finance their spending? Was it ignorance or impotence?

Eric Tymoigne¹ has looked at the written record accumulated over the past two thousand years and found plenty of evidence that the nominalist view was well-represented in monetary debates from the time of Plato (he gives Plato credit as the first to embrace nominalism). It is probably more correct to argue that metalism was often adopted for pragmatic reasons—not due to ignorance. Coining precious metal and promising to convert currency to precious metal offered several advantages over the use of a fiat currency.

First, as discussed in Chapter 3, a sovereign issuer usually could not compel foreigners to accept the currency—the sovereign's ability to impose obligations was limited to their own domain. The advantage of precious metal coins is that they would be accepted externally. Second, the ability to tax even *within* the sovereign's domain was limited—taxpayers could resist, revolt, and evade taxes. Perhaps a strong republic—like that of Rome (or the Greek and Italian city-states)—could impose and collect sufficient taxes to redeem a large issue of currency. The much weaker medieval monarchs—who faced many rivals, including feudal lords and other monarchs—could not do so. As Tymoigne reports:

Most medieval scholars do recognize the possibility of taxation [that, is using taxation to redeem the currency] but see the sovereignty of the King as too weak to apply effectively taxation to the entire Kingdom, whereas feudal lords have more ability to use and manage that power for taxation given the limited size of their

fiefdom. As the taxation ability of the Sovereign and his control over the monetary system grew from the 13/14th century scholars became more willing to consider taxation as a relevant means of redemption. (Tymoigne 2024, p. 41)

Their spending to promote a kingly lifestyle while also conducting wars exceeded their ability to collect taxes. It must also be remembered that sovereigns often relied on “tax farmers” to collect the taxes—which opened the door to fraud through the underreporting of revenue. Thus, sovereigns routinely issued far more currency than they could redeem in taxation.

That led to the third problem: the sovereigns would default on their promises—crying down their coins and defaulting on debt. They might also decide to refuse to accept some kinds of currency (worn or clipped coins) in favor of another (coins over paper notes)—effectively defaulting on a portion of their debt. Issuing precious metal coins provided a degree of insurance to their subjects: the coins would retain some value even after a sovereign default. Relatedly, hired mercenaries wanted precious metals because the crown for whom they fought could lose.

Finally, given the scarcity of precious metals, counterfeiting coins would be more costly than counterfeiting paper notes, for example. While the king’s tally sticks were difficult to counterfeit, they were less useful as a medium of exchange and thus less likely to be counterfeited.

More fundamentally, as Tymoigne argues, there is a difference between the redemption of sovereign currency and the redemption of liabilities issued by banks. The first requires tax collection, implying some measure of coercion and a willingness on the part of the crown to expend the effort to collect taxes. The redemption of bank liabilities, however, occurs naturally to finance commerce. Early banks discounted bills issued to finance trade or made short-term loans to businesses to finance production. Bills were redeemed as trade was completed, and the debtors to banks would redeem banknotes or deposits in payment to get out of debt. They had good reason to redeem the bills, notes, and deposits, and the for-profit banks had good reason to collect the payments. By contrast, recipients of government currency would prefer not to redeem it—no one likes to pay taxes—so if taxes could be avoided or evaded, the currency would not be redeemed. And given the likelihood that the sovereign wanted to spend far in excess of the ability to collect taxes, much of the currency would not be redeemed. This could incentivize the sovereign issuer to devalue the coins or—as in the case of Charles II—refuse his own currency in the form of tally sticks (because he wanted coins). Further, a new king might repudiate the debt of the previous one.

For all these reasons, pragmatic metalism made sense—it limited the sovereign’s ability to spend and protected the holders of the crown’s currency. Indeed, as Tymoigne argues, those who understood and advocated nominalism

also supported austerity—they were by nature fiscal hawks because the crown’s propensity to spend chronically exceeded its ability and willingness to tax.

The default by King Charles set in motion the development of the modern financial system based on nominalism. As discussed in Chapter 3, English banking was not well developed at the time, and bills of exchange continued to play a dominant role. In addition, in England, the interests of a monarch who issued coins and those of producers who issued bills were antagonistic. (Ingham 2004a, p. 203) “At times, from the fourteenth to the mid-17th century, English kings banned the importation of foreign coins and the export of bullion; commanded exporters to supply their bullion to the mints; attempted to prohibit the bill of exchange; and generally sought to limit the use of credit.” (ibid.) All of this delayed the development of the financial system, offering the opportunity for Parliament to chart a new course.

Queen Elizabeth I’s reform (1560–61) established a coinage and prohibited foreign coins, so the money used domestically was coextensive with the state. (Ingham op. cit., p. 205) With Parliament taking the upper hand over the new King William a century later, the “emerging English nation-state became the basis for the impersonal trust that eventually enabled the forms of credit money to become established outside interpersonal banking and exchange networks.” (Ingham op. cit., p. 205) The pound was rigidly tied to silver, and the Bank of England was created to finance the state. Given the underdevelopment of the private banking system, the Bank of England gradually emerged as a central bank that integrated the private and public *giros*: Bank of England notes financed both the public and private spheres—meaning they would be redeemed both in payment of taxes and in payment of private debts. The Bank also financed the government, purchasing its debt and issuing notes against it. With the entire banking system relying on the Bank’s liabilities for clearing (and, eventually, for lender of last resort support), there was always a demand for them.

To be sure, the 19th century saw ample debate over proper monetary management, as discussed in Chapter 5. And the rigid peg of the pound brought its own problems (all the way up to the Great Depression). However, the developments in England that followed the reforms and the creation of the Bank of England, which merged the public and private financial spheres, became the model for the development of the modern financial system—a monetary system well formulated for democratic capitalism.

What is important to understand is that the debates were not about whether the government can spend its currency into circulation, nor about the use of taxes to redeem the currency—these points were well-established, as Tymoigne shows. The question was whether redemption through taxation would be sufficient to maintain the value of the currency. Using precious metal in coins or promising to convert currency to precious metal on demand (the other kind

of redemption) were proposed to further ensure that the currency maintained its value. By the late 19th and early 20th centuries—when Knapp, Innes, and Keynes were writing—these questions were still on the table. As Keynes argued, managed money remained the dominant approach. However, Innes and, later, Lerner, Minsky, and Goodhart came down firmly on the side of fiat money, ending the tie to precious metal.

In Chapter 5, we continue our discussion of credit money, examining the early formation of the endogenous approach to money and briefly contrasting it with the exogenous approach. These debates heated up in the 19th century and continued through the first three-quarters of the 20th century, after the development of the modern monetary system with a central bank at the top. While superficially they appear to be related to the nominalist/metalist debate, the context is actually quite different. Because states no longer spent currency into circulation, these later debates largely took for granted that tax revenues financed government spending, and the role played by taxes in redeeming the currency was largely sidelined. In Chapters 8 and 9, we will examine how modern governments spend and tax through their central banks—the system created in England in the 17th century.

Ironically, the insertion of the central banks and the private banks into the public financial sphere made these operations opaque. Until MMT revived the state money approach, most economists had little understanding of the way a fiat money system operates. It is only recently that orthodox economists have embraced the endogenous money view. And most economists still do not understand how the government really spends.

In Chapters 5 and 6, we look at the endogenous money approaches and will conclude with an integration of credit and state money approaches, before turning to fiscal operations in Chapters 8 and 9.

NOTE

1. Eric Tymoigne, *The Origins of the Platonic Approach to the Analysis of Monetary Systems: Retracing European and Chinese Monetary Thoughts on Chartalism, Nominalism and the Origins of Monetary Systems*, Draft Manuscript, Lewis & Clark College, June 2024.

5. Introduction to the origins of the endogenous money approach

In the last chapter, we distinguished between the Metalist and Chartalist approaches to money. There is a related debate that runs through the history of thought concerning the nature of money—whether it is exogenous or endogenous. While the overlap between the Metalist and exogenous approaches is not complete, both of these approaches are associated with Monetarism—a school of thought that dominated economic theory and policy from the late 1950s through the 1980s. Monetarists recognized that money is no longer linked to metal (although some argued that a return to gold might be ideal if not practical), but they argued that even without a gold standard, it is possible to make the monetary system act as if it were constrained by gold. By contrast, the endogenous money approach insists that private credit money is not under control of the central bank; rather, it expands more or less to finance private activity. What is perhaps more controversial is to claim that there is a natural link between endogenous money and Chartalism—a case we will make in the next chapter.

We begin with a brief overview of the history of the endogenous money approach and then discuss the debate between the currency school and the banking school in 19th-century England as it clearly highlights the differences between the exogenous and endogenous views and is related to our earlier discussion of rival views of the nature of money. We then turn to proponents of the endogenous approach from the second half of the 19th century to WWII. We focus especially on Marx, Keynes, and Schumpeter—three of the giants who emphasized the importance of money in the capitalist economy. In the next chapter, we will look at the postwar revival of the endogenous money approach, largely by Post Keynesians. It is significant that the exogenous money approach has fallen out of favor even among orthodox economists over the past few decades. It is still promoted in the financial press but is no longer accepted by those making monetary policy. While the metalist approach to money is no longer popular, the state money or Chartalist approach is not widely embraced even by those who accept the endogenous money approach.

BRIEF OVERVIEW OF THE HISTORY OF THE ENDOGENOUS MONEY APPROACH

This chapter is not meant to be a comprehensive survey of all those who worked within the endogenous money approach before WWII but to emphasize that the approach is not new. We will see that while the Classical economists admitted that the relation between money and nominal values was a two-way relation, their analysis did not lead to a theory of an endogenous money supply. The Banking School, as represented by Tooke, came much closer to an endogenous approach. Marx adopted an endogenous money view; however, his work has not been of major importance in further development of the approach. Still, it is important to acknowledge Marx's position as we will explore more fully his approach to money in Chapter 10 below. Keynes argued in the *Treatise on Money* and elsewhere that the money supply is endogenously determined but may have assumed that it was exogenously given in much of his exposition in the *General Theory*.

While many economists had adopted an endogenous money approach in the 19th century, the exogenous money view gained adherents in the 1920s with the "discovery" of the so-called deposit multiplier. Ironically, an exogenous money supply was adopted in the neoclassical synthesis (the textbook version of Keynes), and the quantity theory of money was revived after WWII with the development of Monetarism—also with an exogenously determined money supply. Thus, the two major postwar schools (Keynes of the textbooks and Monetarist) had long accepted an exogenous money supply as common ground, so that postwar debates were almost entirely over issues which follow from an exogenously determined money supply (especially over the shape of the LM curve).

We saw in Chapter 4 that Smith recognized that paper money can substitute for gold, and that he and others believed that the redemption of notes reduces the supply of money—which makes the money supply endogenous. Knapp and Innes went further, arguing that although the state chooses the money of account, credit money denominated in that unit is created by banks (Knapp called that money *paracentric*—used as a means of payment between private persons). Innes argued that we should view all sales as really "the exchange of a commodity for a credit" and spoke of the "constant creation of credits and debts, and their extinction by being canceled against one another" as the basis of "the whole mechanism of commerce." (Innes 1913, p. 393) Clearly, those views are consistent with an endogenous money view. But they were alien to postwar orthodoxy that had adopted the view that no individual bank can create money, although with a fractional reserve system they can increase the money supply by a multiple of the reserves supplied by the central bank. They

presented a very truncated view of both money and banking such that—as Friedman claimed—all the important features of capitalism are contained in the barter model of exchange.

In the next chapter, we will examine the modern revival of an endogenous money approach, which can be traced to Kaldor, the Radcliffe Committee, and Minsky. We will then examine the work of Tobin, Gurley and Shaw, and several Post Keynesians, including Moore and Lavoie. In this chapter, however, we will look to the precursors of that revival, focusing on the work of Marx, Keynes, and Schumpeter. Schumpeter, in particular, described money as not only essential to the capitalist economy but also related credit creation to the profit-seeking behavior of capitalists.

THE CURRENCY SCHOOL AND THE BANKING SCHOOL

There are many similarities between the Currency School-Banking School controversy and the debate between Monetarists and Keynesians of the 1960s to the 1980s. The Currency School position is similar to that of Monetarists, while the position of the Banking School is close to that of Keynesians, particularly to that of Post Keynesians. These similarities have been discussed by Cramp (1962, 1970), but he did not present the arguments of the Banking School in detail.

The Currency School included Lord Overstone, Robert Torrens, and George Warde Norman, while the Banking School was represented by Thomas Tooke, John Fullarton, and John Stuart Mill. (Cramp 1962) The debate centered on the relation between money and spending, on the ability of the Bank of England to control the quantity of money, and on the relation between the quantity of money and England's external balance. It took place during the 1830s and 1840s, but before analyzing the debate, we will first examine the views of the Classical economists on money.

Hume analyzed the relation between prices and the money supply and, like other Classical economists, his analysis focused on the effect of a change in prices on the trade balance due to a change in the money supply. The following statement is particularly clear:

Suppose four-fifths of all the money in Great Britain to be annihilated in one night, and the nation reduced to the same condition, with regard to specie as in the reigns of the Harrys and Edwards, what would be the consequence? Must not the price of all labour and commodities sink in proportion, and everything be sold as cheap as they were in those ages? What nation could then dispute with us in any foreign market, or pretend to navigate or to sell manufactures at the same price, which to us would afford sufficient profit? In how little time, therefore, must this bring back the money which we had lost, and raise us to the level of all the neighbouring nations? Where after we have arrived, we immediately lose the advantage of the cheapness

of labour and commodities; and the farther flowing in of money is stopped by our fulness and repletion. Again, suppose, that all the money of Great Britain were multiplied fivefold in a night, must not the contrary effect follow? Must not all labour and commodities rise to such an exorbitant height, that no neighbouring nations could afford to buy from us; while their commodities, on the other hand, become comparatively so cheap, that ... they would be run in upon us, and our money flow out; till we fall to a level with foreigners, and lose that great superiority of riches, which had laid us under such disadvantages? (Hume 1898, p. 333)

Thus, a decline in the money supply causes prices to fall and generates a trade surplus. This causes a specie inflow until the money stock and prices have returned to their original levels. On the other hand, an increase in the money supply raises the price level and results in a trade deficit. Gold flows out of the country until the money supply and price levels are restored to their original levels. The exogenous money supply is linked to gold, and the specie-flow mechanism maintains balanced trade.

Ricardo extended this argument to include paper money. If the quantity of money were to be increased:

[These notes] would be sent into every market, and would everywhere raise the price of commodities, till they were absorbed in the general circulation. It is only during the interval of the issue, and their effect on prices, that we should be sensible of an abundance of money; interest would, during that interval be under its natural level; but as soon as the additional sum of notes or of money became absorbed in the general circulation, the rate of interest would be as high. (Ricardo 1951, Vol. 3, p. 91)

The similarity to the simple monetarist story of the link between money and price is clear: the supply of money determines price, and as the supply rises, prices rise. (As we will see below, while the Banking School rejected that link, they agreed with Ricardo that the interest rate would be impacted.) However, Hume and Ricardo did not offer a detailed analysis of the working of the money market mechanism—there is no explanation of the process through which the money supply is increased or decreased. If the money supply is exogenously increased, this leads to a trade imbalance so that the money supply is restored to its original level. However, there is no theory regarding how the money supply can attain the “correct” level in the first place.

Tooke and J.S. Mill attempted to extend classical theory through an analysis of the role of the banking sector in the economic system. This led Tooke to reject the position of Hume, Ricardo, and most of his contemporaries, who claimed an increase in money directly leads to an increase in the price level. He argued, “the prices of commodities are not liable to be influenced in any perceptible degree by variations in the issues of the Bank, in a convertible state of the paper.”¹ (Tooke 1848, Vol. III, p. 245)

Before this conclusion can be examined, it is first necessary to discuss Tooke's theory of money and to distinguish it from that of the Currency School. Tooke believed that the Currency School reached incorrect conclusions regarding the relation between money and prices because it ignored the essential differences between the ways in which government (fiat) money and private credit enter the economy. The main controversy of the period concerned the proper way to rectify a trade imbalance. The Currency School advocated a tight money policy, while Tooke argued that this disrupts credit markets. Finally, we will examine the monetary policy advocated by Tooke, which included policies to stabilize interest rates and to ensure the solvency of banks.

Tooke and other members of the Banking School distinguished between paper money and paper credit. The term paper money referred to money issued by the government to purchase goods and services. It corresponds to what is now called government fiat money. Paper credit was the term applied to various liabilities issued by banks (such as bank notes) or issued by firms and accepted by banks (such as bills of exchange). The Banking School argued that paper credit is issued on demand to satisfy the needs of depositors as they draw down deposit accounts to purchase consumption goods or to facilitate production by allowing the "circulation of capital":

[A]ll the transactions between dealers and dealers, by which are to be understood all sales from the producer or importer, through all the stages of intermediate processes of manufacture or otherwise to the retail dealer or the exporting merchant, are resolvable into movements or transfers of capital. (Tooke 1959, pp. 35–6)

Bank notes are merely paper credit, as Fullarton argues:

Now, what are bank-notes but a form of credit? What are they but credit parceled out into small and even sums, for greater convenience in circulation? Each note is simply a transferable acknowledgment of a debt due from the banker to the first recipient of the note, and which he (the banker) promises to pay on demand: the value of which it purports to convey is that which it enables the holder to command by sending it to the bank, from whence it issued, and no other ... (Fullarton 1969, pp. 36–7)

Bank notes are issued at the initiative of the depositor and cannot be issued at the will of the bank. Thus, bank notes functioned much as demand deposits function today: issued on demand. Therefore, it is difficult to conceive of an "excessive" supply of bank notes, as discussed below. These positions of the Banking School are in direct contrast to those of the Currency School. Tooke distinguishes between the two schools as follows:

It was held by most writers of any authority on the subject of the Currency, till within the last few years, that the purposes of a mixed circulation of coin and paper

were sufficiently answered, as long as the coin was perfect, and the paper constantly convertible into coin; and that the only evils to be guarded against by regulation, were those attending suspension of payment and insolvency of the banks, a large proportion of which blend an issue of promissory notes with their other business. This, in point of fact, is what is understood in general terms as the banking principle, and is that upon which our system of currency is constructed and conducted. But a new canon of currency has of late been promulgated by persons of no mean authority. According to these authorities, it is not sufficient that the bank notes should be at all times strictly convertible into coin, and that the banks, whether issuing or not issuing, should be solvent; they consider that a purely metallic circulation (excepting only as regards the convenience and economy of paper), is the type of a perfect currency, and contend that the only sound principle of a mixed currency is that by which the bank notes in circulation should be made to conform to the gold, into which they are convertible, not only in value, but in amount. ... A regulation of the issue of bank notes, in conformity with this doctrine, is now understood to be designated as the Currency principle (Tooke 1959, pp. 1–2).

That is, the Currency School wishes to tightly constrain the issue of bank notes, so that the number issued corresponds to the amount of gold coin. The restriction of the issue of bank notes was not primarily to ensure the solvency of banks, but to constrain the amount of paper money issued:

The arguments urged in favour of such separation [of the functions of bank note issue from the functions of a bank as a safe depository] have, as it should seem, made considerable impression on the public mind, and schemes founded upon this principle have been strongly pressed on the attention of government, on the ground not only of guarding against the danger of suspension and insolvencies of the issuers, but of imparting more confidence and stability to credit and trade, and of securing greater steadiness in prices, and thus obviating or abating the alternations of feverish excitement and the extreme of depression, which have prevailed under the existing system, and which are imputed to a neglect of the currency principle (Tooke 1959, pp. 2–3).

Thus, the Currency School, like the Monetarists of recent vintage, argued that control over the supply of money would tame or eliminate the business cycle.

The Banking School-Currency School controversy centered on the correct method of stemming an outflow of bullion. The Currency School accepted the position of Hume and Ricardo that the money supply determines prices. A rise in the supply of money raises prices, causing a trade deficit and an outflow of gold. Thus, the external balance could be restored if the money supply could be controlled. (Cramp 1962, p. 8) Restricting the growth of the money supply would prevent inflation and trade deficits.

At this time, the most important liquid assets included gold coin; “Bank of England and country bank notes; bank deposits subject to cheque; and the ‘promises to pay’ of firms and individuals, which might be embodied in a negotiable instrument (bills of exchange) or might consist in mere book entries

(trade credit).” (Cramp 1962, p. 8). The Banking School emphasized the ability of a wide range of assets to function as money. In particular, transferable bills of exchange could be used to make payments. “It was common practice for B to sell goods to A, draw a bill which A accepted and B then discounted with C, who might use it to settle a debt to his creditor D and so on: it might be G or L or P who eventually presented the bill to A (or his banker) for payment.” (Cramp 1962, p. 10)

Likewise, Mill argued that in the early part of an expansion, trade is financed largely by trade credit. Thus, the relation between Bank of England notes and “near-monies” (see the discussion of Gurley and Shaw in Chapter 6 below), as well as the relation between Bank of England notes and spending, were not believed to be close.

The Currency School, however, included as money only Bank of England notes, specie, and country bank notes. Thus, if issues of these could be controlled, then the quantity of money could be constrained. The Currency School wished to restrain private bank note issues by requiring that banks hold specie or Bank of England notes in an amount equivalent to private bank note issues. Tooke, however, argued that country bank notes are fundamentally different from government note issues (this is the difference between paper credit and paper money alluded to above) and that a country bank note is merely a debt that is not fundamentally different from other private debt issues. In other words, either money should include a wide range of liquid assets, or the term should be limited to the debts of the government plus gold. He argued that it made no sense for the Currency School to include one specific private debt and then to exclude all other private debts from the definition of money.

Tooke and the Banking School argued that private debts should be distinguished from government debt:

[W]hen, in the progress of society, credit comes to perform an important part in all mutual dealings, and in the great majority of transactions supersedes the necessity for this interchange of equivalents, an entirely new principle is brought into play, and one governed by distinct laws. Credit becomes then the legitimate substitute for money, but, in all its modifications and phases, it is distinguished by a broad and impassable line from money itself. The real question, then to be considered is, not whether this or that particular form of credit be entitled to the designation of “money”, but whether, without a perversion of terms and an outrage of principle, that denomination can be applied to credit in any shape. (Fullarton 1969, p. 36)²

Thus, credit substitutes for “money,” but is distinguished from what is now called fiat money. The Currency School, however, confused “Bank notes, strictly convertible into coin, with a compulsory and inconvertible paper currency [fiat money] ... This erroneous impression arises from a neglect of the consideration of the difference in the manner and purpose of the issue.”

(Tooke, 1848, p. 175). Tooke goes on to distinguish government fiat money from private debt issue:

It is quite clear that paper created and so paid away by the government, not being returnable to the issuer, will constitute a fresh source of demand, and must be forced into and permeate all the channels of circulation. Accordingly, every fresh issue beyond the point at which former issues had settled in a certain rise of prices and of wages, and a fall of the exchanges, is soon followed by a further rise of commodities and wages, and a fall of the exchanges; the depreciation being in the ratio of the forcibly increased amount of the issues. It will hence appear that the difference between paper money so issued, and bank notes such as those of this country consists, not only in the limit prescribed by their convertibility to the amount of them, but in the mode of issue. (Tooke, 1848, p. 177)

Thus, government paper money is like Friedman's helicopter money. Tooke simply seems to be arguing that government demand is always excess demand—because its notes are not withdrawn (redeemed, as we discussed earlier). On the other hand, bank liabilities are issued on demand because someone wants to hold or spend them. Bank liabilities do not represent an additional demand because they are merely debt.

The quantity (of bank notes), therefore, is an effect and not a cause of demand. A compulsory government paper, on the other hand, while it is in the course of augmentation, acts directly as an originating cause on prices and incomes, constituting a fresh source of demand in money ... (Tooke 1848, p. 177)

In summary, the Banking School position was that private credit is issued on demand and substitutes for money (narrowly defined as fiat money) to facilitate both the circulation of final goods and the facilitation of production by allowing the circulation of intermediate goods among producers, importers, and retailers. Private credit does not represent an additional monetary demand. In contrast, however, government fiat money (paper money) does represent an additional demand. It differs from private debt issuance because it represents an asset, but not a true liability of the government which ever needs to be retired. Private debt, however, is retired in redemption: bank liabilities are issued to finance relatively short-term assets and will be shortly retired. The quantity of bank liabilities that remain in circulation is decided by the holders of bank debt, and not by banks. Since the quantity of bank liabilities outstanding is demand determined, it can never be excessive. Thus, it can't directly cause prices to rise.

Monetarism, like the Currency School, doesn't distinguish between bank liabilities and fiat money with respect to inflation potential. Both argue an issue of bank notes (or demand deposits after deposit banking becomes dominant) raises prices once holders have decided their money balances are excessive.

However, the Banking School cannot “conceive any means whatever by which the circulation could be so augmented ...” (James Wilson, quoted by Tooke, 1848, p. 195). Furthermore, the Banking School rejects the belief that banks could directly affect the quantity of bank notes outstanding, as this would be determined by the rate of “reflux” (what we have called redemption), which is determined by the holders of bank notes. Tooke describes his *law of reflux*:

This law operates in bringing back to the issuing banks the amount of their notes, whatever it may be, that is not wanted for the purposes which they are required to serve. The reflux takes place chiefly in two ways: by payment of the redundant amount to a banker on a deposit account, or by the return of notes in discharge of securities on which advances have been made. A third way is that of a return of the notes to the issuing bank by a demand for coin. (Tooke, 1848, p. 186)

His law of reflux helps to further distinguish private debt issues from government fiat money. Private institutions are limited in the quantity of debt they can issue due to reflux, while the government is not “because there is no reflux.” (ibid.) However, if private institutions issue more debt than the public needs to facilitate circulation, reflux will increase. If the needs of circulation increase, private debts will be issued to meet these needs. Thus, control over banknotes alone would not control private debt issues, as other forms of debt would be issued in their place to facilitate circulation.

The Currency School responded with the argument that trade credit must finally be retired using gold or Bank of England notes; thus, a stable proportion must exist between *near-monies* and money, narrowly defined. Therefore, “to restrict the flow of monetary demand, it was necessary to reduce the whole stock of liquid assets, and this could be done by squeezing the money supply [narrowly] defined.” (Cramp 1962, p. 11) In Tooke’s view, this attempt would neither affect the quantity of credit used to support circulation nor result in falling prices. If the Bank of England reduced its note issues, these would have to be made up by an expanded circulation of private issues of notes and bills. If the quantity of private issues were not increased, then notes and gold previously deposited in banks would be withdrawn to meet the needs of circulation. In this case, the quantity of money available as “capital” (deposits which could be intermediated by banks) would fall so that the quantity of money available as means of circulation could be restored. This would increase the rate of interest, leading to gold inflows. Equilibrium in the external account would be restored when the interest rate was reduced to its previous level and the money supply was increased to meet the needs of circulation and production.

Thus, the difference between the two schools concerns the mechanism through which “tight money policy” works: the effect is through prices according to the Currency School, while it is indirect and through the interest rate according to the Banking School. The Currency School is clearly the precursor

of the modern quantity theory of money: tight monetary policy directly reduces prices. The Banking School position is a forerunner of Keynesian theory: monetary policy works indirectly through the interest rate. The Currency School advocated tight control over bank note issuance in the belief that this would stabilize the economy by stabilizing prices. Tooke, however, argued that tight monetary policy actually works through destabilizing credit markets.

To understand how policy works, we need to return to an examination of the external account and the proper way to rectify a trade deficit. The Currency School believed that constriction of the money supply (bank notes) would lower prices and restore a favorable trade balance. The Banking School, however, argued that constriction of paper credit couldn't directly affect prices (although constriction of paper *fiat* money could). Tight monetary policy would stem a trade deficit only by raising interest rates. Tooke argued that this policy would disrupt credit markets because large swings in interest rates would be necessary to rectify trade imbalances.

However, the Currency School was able to prevail with policymakers, and tight monetary policy was adopted in the Peel Banking Act of 1844 that severely limited the issuance of bank notes. Tooke argued:

As the result of a careful examination of the principle on which the Act of 1844 was founded, and of the experience of its working since the time when it came into operation, I have no hesitation in giving it as my opinion that it is a *total unmitigated, uncompensated*, and, in its consequences, a *lamentable failure*. (Tooke 1848, p. 402)

The variations in the rate of interest, during the year 1847, have exceeded in frequency and extent any of which there is to be found an example in the commercial history of this country. (op. cit., p. 400)

The policy “worked” by affecting interest rates (as Ricardo had argued), not the money supply broadly defined. The alternative policy advocated by Tooke included “a total abrogation of those provisions of the Acts of 1844 and 1845, which limit the amount of note circulation.” (Tooke 1848, p. 402) Rather than using tight monetary policy, Tooke advocated the use by the Treasury of bullion outflows whenever a trade deficit occurred. Of course, this would only be a temporary measure, but he believed that it would allow lower volatility of interest rates. In order to free bullion for this purpose, it was necessary to allow the expansion of bank notes (and other private debt issues) to meet the needs of circulation. Therefore, restrictions on bank note issues had to be removed to free bullion for use in the external account.

Finally, the proper role for the central bank, according to the Banking School, is to maintain the solvency of banks and stable credit market conditions, Tooke argued:

[I]t is only through the rate of interest and the state of credit that the Bank of England can exercise a direct influence on the foreign exchanges ... [T]he greater or less liability to variation in the rate of interest constitutes, in the next degree only to the preservation of the convertibility of the paper and the solvency of banks, the most important consideration in the regulation of our banking system. (Tooke 1959, p. 124)³

These certainly sound like “Keynesian” propositions (see Chapter 7). The central bank should abandon quantity constraints and concern itself with interest rates and bank solvency. The Banking School is also ancestrally related to the endogenous money approach because it focuses on the use of credit to finance expenditure. In particular, the Banking School, like the endogenous money approach, rejects the simple notion that tight money policy would directly decrease the quantity of money (broadly defined), and rejects the notion that a decline in the quantity of money would lower prices. According to the Banking School, the quantity of money privately issued can never be “excessive” because the needs of circulation determine how much money circulates. Any quantity of private debt issued above this would increase “reflux” (or redemption) and thereby destroy money. Tight policy operates indirectly through interest rates, and, by disrupting credit markets, through aggregate demand.

MARX AND ENDOGENOUS MONEY

Marx analyzed the Banking School-Currency School controversy in detail. He rejected the theories of the Currency School, such as the notion that an increase in the money supply increases prices. However, he accepted some aspects of the theory of the Banking School, which he considered the top monetary theory of that period. He also criticized the Banking School because of the absence of a theory of value (and surplus value) in the approach. Much of his criticism is not directly relevant to the endogenous money approach, so here we will only briefly examine Marx’s system and his critique of the Banking School. We will return to a discussion of Marx’s approach to money in Chapter 10.

Marx’s generalized system can be depicted as $m - m - c \dots p \dots c' - m' - m''$. In the first step ($m - m$), the money capitalist alienates money by lending it to the industrial capitalist. In the second step ($m - c$), the industrial capitalist purchases commodities (for example, labor power and raw materials), which are used in the third step ($c \dots p \dots c'$) to produce commodities (c') with different use values for sale. In the fourth step ($c' - m'$), the industrial capitalist sells the produced commodities, and where $m' > m$, the industrial capitalist can repay the loan from the money capitalist. For interest to be paid, it is necessary that $m'' > m$. Thus, m'' is paid to the money capitalist, so that if $m' > m''$, the

difference is profit to the industrial capitalist. (Thus, m' less m'' is received by the industrial capitalist as profit, while m'' less m is interest received by the money capitalist.)

The money capitalist will only loan money (m) if it is expected that the loan will be repaid with interest ($m' > m$). This requires that produced commodities (c') be sold at a price high enough to enable the industrial capitalist to repay the loan with interest. The industrial capitalist will not undertake the project unless the net profit (m' less m'') is expected to be sufficiently high. Production is undertaken to realize profits through producing goods and services for sale. Ultimately, profits must be realized in money form, rather than in stocks of unsold goods. Thus, capitalist production necessarily takes the form of monetary production: the capitalist must produce goods for the market to earn money to retire debt and leave profit in money terms.

Following Marx, Foley (1989) argues that in capitalism, money measures the value of commodities as a universal equivalent. (See Chapter 10.) It also functions as a unit of account—as the unit in which promises to pay are measured. Financing determines who will be able to carry out planned expenditures. An entity that desires to undertake spending greater than income issues its own liability, which entails a promise to pay later in the form of a third-party liability that is higher in the *debt pyramid*. That is, one's own liability is issued as one goes into debt, but normally the debt can only be retired through the use of another's liability—usually higher in the debt pyramid. Individuals retire debts by using bank liabilities, while banks retire debts by using government (usually the central bank) liabilities. The debtor must operate in the market to obtain these third-party liabilities so that debt, whose value is denominated in the universal equivalent (money), may be retired. Capitalist accumulation (which takes the form of the accumulation of the means to further accumulation) requires that profits remain after debt is serviced. Finally, successful accumulation is measured in terms of the universal equivalent, money.

Marx believed that the supply of credit normally expands to meet the needs of economic activity. He noted that a large portion of loans is created for the purpose of transforming capital in physical form (including inventories) into liquid form until the physical capital can be liquidated (through depreciation or sale) and the credit retired.

The ordinary businessman discounts, in order to anticipate the money-form of his capital and thereby to keep his process of reproduction in flow; not in order to expand his business or secure additional capital, but in order to balance the credit he gives by the credit he receives. (Marx 1972b, p. 525)

He argues that most of the credit created is not used in final purchases but in intermediate stages, quoting Coquelin affirmatively:

In every country the majority of the credit transactions takes place in the circle of the industrial relations themselves ... the producer of the raw material advances it to the capitalist, who works it up, and receives from him a promise to pay on a certain day. The manufacturer, having completed his share of the work, in his turn advances his product on similar conditions to another manufacturer, who has to manipulate it farther, and in this way credit extends more and more, from one to the other, down to the consumer ... All borrow with one hand and lend with the other, sometimes money, but more frequently products. In this manner an incessant exchange of credits, combining and crossing in all directions, takes place in the industrial relations. The development of credit consists precisely in the multiplication and growth of these mutual credits, and here is the real seat of its power. (Coquelin, from *Du Credit et des Banques dans L'Industrie*, 1842, quoted in Marx 1909, p. 472)

Thus, most credit takes the form of working capital. In the absence of credit, the industrial capitalist would be constrained by the quantity of money hoards (either their own funds or those of monetary capitalists). Marx argued that credit is the creation of industrial capitalists to free them from the necessity of accumulating money hoards or from borrowing from those who do. The creation of financial institutions allows the concentration of money capital that can be used as the basis of credit extension:

Production on a large scale and for distant markets throws the total product into the hands of commerce; but it is impossible, that the capital of a nation should be doubled in such a way, that commerce by itself would be able to buy up the entire national product with its own capital and to sell it again. Credit is, therefore, indispensable here. Credit must grow with the growing volume of value in production, and it must grow in the matter of time with the increasing distance of the markets. A mutual interaction takes place here. The development of the process of production extends the credit, and credit leads to an extension of industrial and commercial operations. Looking upon this credit separate from banking credit, it is evident that it grows with an increasing volume of industrial capital itself. (Marx 1909, pp. 564–5)

The creation of credit allows a small amount of money capital to support a large circulation because money capital is only required when credit must be settled, and only for the amount of net credits. As Marx put it:

It is here indisputably true that the 1,000 p.st. [pounds sterling] which I deposit to-day with A are issued to-morrow and deposited with B. The day after to-morrow it may be issued once more by B and form a deposit with C, and so forth infinitely. The same 1,000 p.st. of money may, therefore, multiply themselves into an absolutely indeterminable sum of deposits by a series of transfers. Hence it is possible that nine-tenths of all deposits in England may have no other existence but that in the entries of the banker's books, of whom every one stands good for his part of them. (Marx 1909, p. 478)

Bankers are thus able to expand loans to a multiple of their reserves of money capital. The size of this multiple depends on the degree of development and organization of the banking system. Furthermore, the willingness of banks to extend loans also depends on their confidence in the future. Marx “stressed banks’ promptness to modify their money-reserve—either by refusing to make new advances or by giving bills drawn on the bank itself, instead of money, in payment to their customer—as soon as the degree of uncertainty as to future events increased.” (Panico 1982, p. 70) Thus, the size of the “deposit multiplier” was variable (and potentially limitless).

Once credit has been created, it can function like money as a medium of exchange.

Credit-money springs directly out of the function of money as a means of payment. Certificates of the debts owing for the purchased commodities circulate for the purpose of transferring those debts to others. On the other hand, to the same extent as the system of credit is extended, so is the function of money as a means of payment. (Marx 1978, p. 139)

With the development of commerce and of the capitalist mode of production, which has an eye only to the circulation, this natural basis of the credit system is extended, generalised, elaborated. Money serves here on the whole merely as a means of payment, that is to say, commodities are not sold for money, but for a written promise to pay for them at a certain date. We may comprise all these promises to pay for brevity’s sake under the general category of bills of exchange ... To the extent that they ultimately balance one another by the compensation of credits and debts, they serve absolutely as money, since no transformation into actual money takes place ... These do not rest on the circulation of money ... (Marx 1909, p. 470)

Regarding the role of banks and bank notes, Marx said:

The loan is made (we refer here only to the commercial credit in the strict meaning of the term) by discounting bills of exchange, that is, by converting them into money before they come due, and by advances in various forms ... The credit given by a banker may assume various forms, for instance, that of exchanges on other banks, checks on them, opening of credit in the same way, finally, in the case of banks entitled to issue notes, the bank notes of the bank itself. A bank note is nothing but a draft upon the banker, payable at any time to the bearer, and substituted by the banker for private drafts. This last form of credit appears particularly important and striking to the layman, first, because this form of credit money steps from the mere commercial circulation into the general circulation and serves as money there ... (Marx 1909, p. 474)

Thus, like Tooke, Marx argues that bank notes are merely another private debt, but one which appears to be different from other private debts because of the degree to which it circulates. Marx disagrees with the position of Tooke that banknotes cannot represent an additional demand on goods and services; thus,

he rejects the strict dichotomy Tooke posited between paper credit and paper money. Marx argues that a bank loan may lead to an increased demand, just as fiat money would in the scheme laid out by Tooke: “the bank may open a credit account for A, in which case this A, a debtor of the bank, appears in the role of an imaginary depositor.” (Marx 1909, p. 540) Thus, the borrower receives an “imaginary” deposit that can be spent just as if it were money.

Tooke’s mistake lies in the implicit assumption that all credit advanced by banks is on the basis of existing “capital”—that is, merely to facilitate circulation. However, Tooke was correct in emphasizing that bank credits can rise only in response to a demand for bank credit. In contrast to Tooke, Marx emphasizes that the quantity of notes issued need not be strictly based upon existing “capital,” so can represent an additional demand. Tooke emphasized only the role credit plays in circulating goods that have already been produced, while Marx also emphasized the role credit plays in allowing production to increase.

According to Marx, the quantity of money is a function of the price level, velocity, and the quantity of transactions. All these determinants of the quantity of money are variable. He argues that the theory that an increase in the quantity of money necessarily causes prices to rise is wrong, as the quantity of money tends to increase or decrease with the needs of circulation. He says, “We have seen how, along with the continual fluctuations in the extent and rapidity of the circulation of commodities and in their prices, the quantity of money current unceasingly ebbs and flows.” (Marx 1978, p. 134) He argues:

The degree of the velocity of circulation, in other words, the number of repetitions of the same function as means of purchase and payment by the same pieces of money in a given period of time, the mass of simultaneous purchases and sales, or payments, the sum of the prices of the circulating commodities, finally the balances of payments to be spared in the same period, determine in either case the mass of the circulating money, of currency. Whether the money so serving represents capital or revenue for the paying or receiving party, is immaterial, and does not alter the matter in any way. Its mass is simply determined by its function as a medium of purchase and payment. (Marx 1909, p. 527)

The quantity of money is not a cause, but an effect, of spending. Attempting to control prices by reducing the quantity of money will not be effective because the quantity of money only indirectly affects prices through the cost of production. He says, “in fact the reduction of the quantity of gold raises only the rate of interest, whereas an increase in the quantity of gold lowers the rate of interest; and were it not for the fact that the fluctuations of the rate of interest are taken into account in the determination of cost-prices, or in the determination of demand and supply, the prices of commodities would be wholly unaffected by them.” (Marx 1909, p. 648) Like Keynes and Tooke, Marx argues

that changes in the quantity of money (even narrowly defined) only indirectly affect prices through changes in interest rates.

Marx argued that the demand for credit could be highest during a crisis. In a prosperity phase, capitalists can sell their commodities on the basis of bills of exchange. During the expansion, sellers are willing to accept these bills since their liquidity is high. However, during a downturn, holders of bills attempt to discount them, so the demand for money capital (narrow money) rises. According to Marx, "In time of crisis, the demand for loan-capital, and therefore the rate of interest, reaches its maximum". (Marx 1972a, p. 513). "In times of crisis, loans are taken up for the purpose of paying, not of buying; for the purpose of winding up previous transactions, not of beginning new ones." (Marx 1909, p. 542).

Furthermore, in a crisis, "the circulation of notes as a means of purchase is decreasing" even though "their circulation as means of payment may increase... [and] the aggregate amount of the circulation ... may remain stable or may even decrease." (Marx 1909, p. 542). In other words, over the course of a crisis, the aggregate quantity of debts falls as they are retired and because fewer new debts are taken out. He argues that although the demand for credit will be high in a crisis, that demand won't necessarily be met. In contrast, high demand for credit in an expansion will be met. Thus, high demand for credit cannot be taken as an indication of high expectations: "It is by no means the strong demand for loans ... which distinguishes the period of depression from that of prosperity, but the ease with which this demand is satisfied in periods of prosperity, and the difficulties which it meets after a depression has become a fact." (Marx 1909, p. 532).

Marx's position shares several characteristics with the endogenous money position. During an expansion, the basis of most expenditure is credit expansion. The development of financial institutions reduces the cash reserves that must be held against debt commitments. Debt instruments are normally sufficiently liquid to enable circulation to proceed. In an expansion, credit money expands to meet the needs of circulation. The money supply is not exogenously determined and does not limit production and circulation, since various credit arrangements allow the creation of near monies. There is no fixed ratio of near monies to reserves (cash, or high-powered money in modern parlance). In an expansion, minimal reserves are required to maintain the liquidity of credit instruments such as bills of exchange, while in a crisis large reserves may be required to meet attempts at liquidation. Because an expansion is based on an increase in the quantity of credit relative to cash, it is impossible to exchange all credit for cash during a downturn. Hence, "It is precisely the enormous development of the credit system during a period of prosperity, hence also the enormous development of the demand for loan capital and the readiness with which the supply meets it in such periods, which brings about a shortage of

credit during the period of depression.” (Marx 1909, p. 532) The money supply (and types of credit instruments that function as money) varies endogenously with the business cycle.

KEYNES AND ENDOGENOUS MONEY

Keynes’s conception of money is consistent with the endogenous approach adopted here. In his *General Theory*, he seems to adopt an endogenous approach about as often as he adopts an exogenous approach to money, but that was probably to simplify the analysis. Especially in Chapters 13 and 15, he provides a formulation that led to the framework adopted by Hicks’s IS-LM analysis, with a fixed money supply and what appears to be a downward-sloping money demand curve. It has been argued (Foster 1986a, 1986b) that his adoption of an exogenous money supply seriously weakened the arguments in the *General Theory*, and some Post-Keynesians have objected to his liquidity preference approach to interest rates on the argument that this, too, adopts an exogenous money supply. We will revisit this criticism later in Chapter 7. Elsewhere in the *General Theory*, and especially in his *Treatise on Money*, Keynes clearly adopts an endogenous money approach. That is what we will focus on here.

Above, we have emphasized the role of money as a unit of account, as did Keynes in his writings on ancient currency:

An article be deemed to have some at least of the peculiar characteristics of money 1) if it is regularly used to express certain conventional estimates of value such as religious dues, penalties or prizes, or 2) if it is used as the term in which loans and contracts are expressed, or 3) if it is used as the term in which prices are expressed, or 4) if it is used as the habitual medium of exchange. In the first three cases the article in question is the term in a money-of-account, in the fourth case it is used as actual money. Now for most important social and economic purposes what matters is the money of account; for it is the money of account which is the subject. (Keynes 1982, p. 252)

In the *Treatise*, Keynes said: “Money proper in the full sense of the term can only exist in relation to a money of account ...” (Keynes 1971b, p. 3) He goes on to distinguish among various forms of money which fit the description—as discussed in Chapter 4. Money “proper” is defined as state money plus those private forms of money that are accepted by the state. Thus, as discussed earlier, when the state becomes a member of the bank *giro* by accepting payment of bank notes or deposits, bank money achieves a special status, which Keynes calls money proper. Other forms of money that fulfill the unit of account function are merely “offers of contracts, contracts and acknowledgments of debt,” which exist “side by side” with money proper. (ibid., p. 5) Furthermore,

“for many purposes the acknowledgments of debt are themselves a serviceable substitute for money proper in the settlement of transactions.” (ibid., p. 5) Keynes later criticized the orthodox view of money as a medium of exchange by emphasizing the second major role of money: protection against an uncertain future.

Why should anyone outside a lunatic asylum wish to use money as a store of wealth? Because, partly on reasonable and partly on instinctive grounds, our desire to hold money as a store of wealth is a barometer of the degree of our distrust of our own calculations and conventions concerning the future ... The possession of actual money lulls our disquietude ... (Keynes 1973, pp. 116–7)

Money has real effects because the premium required to induce people to part with liquidity affects the price of capital assets, which then affects the level of employment and output.

(T)he rate of interest is the reward for parting with liquidity for a specified period ... It is the “price” which equilibrates the desire to hold wealth in the form of cash with the available quantity of cash ... (Keynes 1964, p. 167)

[I]t is the greatest of the own-rates of interest ... which rules the roost (because it is the greatest of these rates that the marginal efficiency of a capital-asset must attain if it is to be newly produced); and ... there are reasons why it is the money-rate of interest which is often the greatest. ... (op. cit., pp. 223–4)

Thus, money plays an important role in Keynes’s theory, where “the firm is dealing throughout in terms of sums of money. It has no object in the world except to end up with more money than it started with.” (Keynes 1979, p. 89) He calls his theory a “Theory of a Monetary Economy”, which is: “the theory of a system in which changing views about the future are capable of influencing the present situation. For the importance of money essentially flows from its being a link between the present and the future.” (Keynes 1964, p. 293) Optimism about the future lowers the liquidity premium, allowing production to proceed. A given quantity of money “provides a revolving fund for a steady flow of activity; but an increased rate of flow needs an increased stock to keep the channels filled.” (Keynes 1973, p. 230) Because a reduction in liquidity preference lowers the premium required to release idle liquid balances, production can expand initially with no increase in the money stock. (We revisit liquidity preference theory in Chapter 7.)

Once idle balances have been released, the limit to the rise in production depends on the quantity of finance, the propensity to hoard, and the velocity of circulation. Keynes argues that spending can be constrained by a shortage of finance, but never by a shortage of saving. (Again, see Chapter 7.) In a monetary economy, saving is not equivalent to finance.

Saving has no special efficacy, as compared with consumption in releasing cash and restoring liquidity ... Consumption does just as well; in fact it does better since an increase in wealth may involve a more or less proportionate increase in the inactive demand for cash. A given level of activity and income will involve the same active demand for cash, if the technical conditions governing the time lags are the same, irrespective of the current rate of net investment and saving - irrespective, indeed of whether there is any investment and saving ... There is, therefore, just as much reason for adding current consumption to the rate of increase of new bank money in reckoning the flow of cash becoming available to provide new 'finance', as there is for adding current saving. (Keynes 1973, p. 233)

In other words, consumption is as much a part of loanable funds as is saving if funds spent on consumption find their way into the banking system. Furthermore, since all expenditures are received as income, and because investment goods are unavailable for consumption, that part of income not spent on consumption goods must be equal to expenditures on non-consumables. The supply and demand of "loanable funds" are necessarily identical, with no adjustment in the interest rate required:

Increased investment will always be accompanied by increased saving, but it can never be preceded by it. Dishoarding and credit expansion provides not an alternative to increased saving, but a necessary preparation for it. It is the parent, not the twin, of increased saving. (Keynes 1973, p. 281)

Keynes goes on to argue that while investment can never be constrained by a lack of saving, it can be constrained by lack of finance. He argues that firms may attempt to accumulate a cash balance before undertaking investment. He says:

[U]nless the banking system is prepared to augment the supply of money, lack of finance may prove an important obstacle to more than a certain amount of investment decisions being on the tapis at the same time. But "finance" has nothing to do with saving. At the "financial" stage of the proceedings no net saving has taken place on anyone's part, just as there has been no net investment. (Keynes 1937, p. 247)

So, by the *finance motive*, Keynes means a propensity to build a hoard of cash in preparation for funding an investment project. This must be added to other motives for hoarding. The intersection of the demand for hoards and the supply of hoards then determines the interest rate. A rise in investment doesn't directly affect the interest rate, which is determined by the propensity to hoard. However, if firms accumulate a hoard in preparation for investing, this does cause interest rates to rise unless someone meets this additional demand for a hoard.

There has, therefore, to be a technique to bridge this gap between the time when the decision to invest is taken and the time when the correlative investment and saving actually occur. This service may be provided either by the new issue market or by the banks ... If investment is proceeding at a steady rate, the finance (or the commitments to finance) required can be supplied from a revolving fund of a more or less constant amount, one entrepreneur having his finance replenished for the purpose of a projected investment as another exhausts his on paying for his completed investment. But if decisions to invest are (e.g.) increasing, the extra finance involved will constitute an additional demand for money ... (Keynes 1973, pp. 208–09)

To summarize, an increase in investment can be funded by credit (or releases of hoards). As investment occurs, this creates saving. Since most credit advanced by banks to fund investment spending returns to banks (the “revolving fund”), no pressure is placed on interest rates. However, if part of the credit advanced by banks ends up as hoards (for example, to satisfy the finance motive), then investment can be constrained by lack of finance. Either banks will have to provide this extra money that is hoarded, or interest rates will rise to decrease the desire to hoard sufficiently so that the existing quantity of hoards satisfies the desire to hoard.

There may be some confusion or oversimplification in Keynes’s presentation. He correctly argues that saving is a result of investment, so it cannot be a net source of funds available to finance investment. He also correctly argues that the interest rate cannot be determined by the intersection of the saving supply and investment demand curves. He argues that it is liquidity preference, simplified as the demand and supply of hoards, that determines interest rates. When investment is proceeding at a steady pace, a revolving fund of a given size is adequate to finance that level of investment, and investment generates an equivalent amount of saving. When investment increases, it may place pressure on interest rates—not because it is necessary to increase the quantity of funds available to finance investment but to provide the hoards necessary to satisfy a “finance motive.” He argues that because investment, itself, can be financed out of a revolving fund, it does not place pressure on interest rates. Only money “sitting still” (hoards) affects interest rates; money “on the wing” does not. All of this seems fine.

However, a problem with his analysis arises because he has not adequately explained how the rise of investment is funded. In his theory, if investment were to increase without requiring an increase in the hoards necessary to satisfy the “finance motive,” then it would be funded with no impact on interest rates. However, Keynes is not clear whether the increase in investment would have any impact on the money supply. This is why some Post Keynesian critics argue that he seems to have adopted an exogenous money approach.

But Keynes argued many times that the money supply is not exogenously fixed. For example, he argued:

The banking system has no ... direct control over the quantity of money; for it is characteristic of modern systems that the central bank is ready to buy for money at a stipulated rate of discount any quantity of securities of certain approved types. (Keynes 1971c, p. 189)

The volume of the paper money, on the other hand, would be consequential, as it is at present, on the state of trade and employment, bank-rate policy and the Treasury bill policy. (1971a, p. 153)

Note that by “banking system” Keynes is including the central bank—so he is implicitly rejecting the view that central banks can control the amount of bank money created by the “system” as in the deposit multiplier story.

In the *Treatise on Money*, he recognized that banks create money as they make loans: a bank “may itself purchase assets, i.e. add to its investments, and pay for them, in the first instance at least, by establishing a claim against itself.” (Keynes 1971b, p. 21) Although banks would then lose reserves during a clearing drain, Keynes recognized that “it is evident that there is no limit to the amount of bank money which the banks can safely create provided that they move forward in step,” (ibid., p. 23) since losses through the clearing drain would be compensated by gains. However, he concluded that “the rate at which an individual bank creates on its own initiative is subject to rules and limitations” such as the rate of growth of other banks, rules of thumb and laws concerning reserve ratios, and “the aggregate of their reserve resources.” (ibid., pp. 26–7) Thus, the money supply depends on a number of variables and should not be taken as exogenously fixed.

Furthermore, in a very interesting passage, Keynes examined “overdraft facilities,” in which credit is, by nature, advanced on demand.

[T]o the extent that the overdraft system is employed and unused overdrafts ignored by the banking system, there is no superimposed pressure resulting from planned activity over and above the pressure resulting from actual activity. In this event the transition from a lower to a higher scale of activity may be accomplished with less pressure on the demand for liquidity and the rate of interest. (Keynes 1973, pp. 222–3)

In other words, the existence of overdraft facilities reduces the hoards desired to satisfy the finance motive and automatically increases the supply of money to hoard as long as the banking system accommodates. This, in turn, allows investment to proceed with little impact on interest rates. If Keynes had pursued his analysis of overdraft facilities, he might have more explicitly adopted an endogenous money approach in the *General Theory*, in which money is supplied on demand. Indeed, he later argued that banks provide the short-term finance necessary to allow investment to proceed. As investment generates

saving, an increase in investment can be permanently funded out of the resulting saving.

[T]he increment of current investment over prior investment (or saving) can only be cared for permanently out of the increment of current saving; and the period during which current savings are kept liquid by their owners must be bridged by an increase in the revolving fund of “finance” ... It is the role of the credit system to provide the liquid funds which are required first of all by the entrepreneur during the period before his actual expenditure, and then by the recipients of this expenditure during the period before they have decided how to employ it. (Keynes 1973, pp. 284–5)⁴

While this is a clear statement of endogenous money, Keynes sometimes seemed to assume an exogenous money supply in the *General Theory*. For example, he writes, “the quantity of money as determined by the action of the central bank,” (Keynes 1964, p. 247) “the quantity of money created by the monetary authority,” (op. cit., p. 205) and makes similar arguments in other places. (See also Keynes 1964, pp. 84, 167, 174, 230, 267.) On the other hand, it seems that this assumption was made purely to facilitate the analysis. For example, he argues, “we can draw the line between ‘money’ and ‘debts’ at whatever point is most convenient for handling a particular problem.” (Keynes 1964, p. 167) He then goes on to say that he will arbitrarily define money as bank deposits. If money were defined to include a wider range of debts, then it would be clear that Keynes would not have accepted an exogenously determined money supply.⁵

Joan Robinson argues that Keynes

had to make every possible concession ... in order to get a hearing. It would have been much simpler to start by assuming a constant rate of interest and a perfectly elastic supply of money. But then his whole case would have been dismissed as a misunderstanding of the orthodox position. He was obliged to accept the presumptions of his critics in order to explode them from within. (Robinson 1971, pp. 81–2)

Thus, it is possible that Keynes sometimes adopted an exogenous money supply to facilitate the exposition of various components of his *General Theory*. Most importantly, in many—maybe most—of the passages that seem to adopt an exogenously fixed money supply, he is discussing the possible impacts of an increase in the demand for money. In other words, he holds the money supply constant and analyzes what would happen to interest rates if money demand were to rise. It is simply the usual “demand and supply” approach in which one is held constant while the other changes. And this is most frequently done in Chapters 13 and 15 of the *General Theory*. He then presents a *general theory* of money, liquidity preference, and asset pricing in Chapter 17. The earlier chapters are a simplified—and ultimately misleading—version

that equates “liquidity preference theory” with a “money supply and demand” theory. Chapter 7 below will pursue the more general approach.

Much of his other work seems to recognize that the money supply is endogenously determined, and even the *General Theory* might have been strengthened had Keynes consistently used an endogenous money supply. In Chapter 7, we will see how liquidity preference theory strengthens the endogenous money approach. Indeed, endogenous money combined with Keynes’s theories of effective demand and liquidity preference together provide the basis for a Post Keynesian general theory of money—or, of the *monetary theory of production*.

SCHUMPETER ON CREDIT AND INNOVATION

Schumpeter offered a unique vision of money, arguing that credit is necessary for economic development because it gives purchasing power to entrepreneurs, enabling them to foster new employment of labor and other resources. Economic change in a capitalist society requires that entrepreneurs innovate, and credit allows them to embark on new enterprises. While an existing supply of money can circulate a constant amount of capital, innovation requires new purchasing power that must be financed.

He argues that economic development consists of finding new ways to combine labor and natural resources. Great strides in development do not come from “adaptive responses” but from “creative responses,” which escape current behavior patterns. The “creative response changes social and economic situations for good.” (Schumpeter 1951, p. 217) In a capitalist society, as opposed to a command or communal society, “the mechanisms of economic change ... pivot on entrepreneurial activity.” (ibid., p. 217) The drive for above-normal profits leads to innovation, which generates “capital gain, that is, ... the discounted value of the stream of prospective returns.” (op. cit., p. 222) The capitalists that innovate gain a competitive advantage, while “the impact of the new product or method spells losses to the ‘old’ firms. The competition of the man with a significantly lower cost curve is, in fact, the really effective competition that in the end revolutionizes the industry.” (op. cit., p. 223)

In order to innovate, entrepreneurs must have purchasing power so that they can employ labor and resources. Schumpeter argues that credit is not necessary to maintain a circular flow in equilibrium: the production of consumption goods just satisfies the demand for consumption goods, which is determined by the portion of income spent on consumption, while the portion of income not consumed just equals expenditures on the means of production required to produce the consumption goods. However, credit is required to enable production to increase, as the income arising from a circular flow in equilibrium

cannot be large enough to enable spending to rise. Credit involves the creation of new purchasing power. (Schumpeter 1934, Chapter 3)

In a capitalist economy, “credit is essentially the creation of purchasing power for the purpose of transferring it to the entrepreneur.” (Schumpeter 1934, p. 107) The creation of this purchasing power is “the method by which development is carried out in a system with private property and division of labor.” (ibid.) Credit transforms the entrepreneurs into debtors to society: they receive resources and labor from society before they have contributed to society. They repay this debt to society only after engaging in production. Schumpeter defines capitalism as a form of economic organization in which “the goods necessary for new production” are withdrawn from the circular flow by the intervention of purchasing power “created *ad hoc*” through the creation of credit. (op. cit., p. 106) Credit is capital, which is “nothing but the lever by which the entrepreneur subjects to his control the concrete goods which he needs, nothing but a means of diverting the factors of production to new uses, or of dictating a new direction to production.” (op. cit., p. 116)

When entrepreneurs use credit to purchase productive means, they surrender “capital” as it becomes income to those from whom they purchased the means of production. Capital is “that sum of means of payment which is available at any moment for transference to entrepreneurs.” (Schumpeter 1934, p. 122) The means of payment are not merely “metal money” but include claims to money that take the place of money (narrowly defined). Banks create money by issuing claims against themselves in the process of granting credit. The aggregate quantity of outstanding credit greatly exceeds the quantity of “metal money.” The “circulating media ... exist in such quantities that they could not possibly all be redeemed at once ...” (op. cit., p. 98) Credit is created to give purchasing power to the entrepreneur, but that credit then “increases the means of payment” as it circulates. (op. cit., p. 99)

When entrepreneurs go to a bank for credit, they can offer securities or the goods they will produce as collateral. In the second case, the credit comes before the collateral exists. Thus, the quantity of outstanding credit exceeds the quantity that is “fully covered” by securities or existing goods functioning as collateral. The limit on credit expansion by banks is determined by the necessity that banks make profits. Banks are subject to three constraints: first, they must hold reserves in case the expected production of goods does not occur and the entrepreneur defaults on debt (a loan-loss reserve); second, banks cannot create coined money for transactions, so they must maintain a reserve to meet the demand for currency (liquidity); and third, credit inflation can lead to an outflow of gold. (Schumpeter 1934, pp. 112–4) This last point will require some explanation of *credit inflation*.

Because credit creation gives new purchasing power before new goods are produced, it necessarily generates credit inflation.⁶

Just as when additional gas streams into a vessel the share of space occupied by each molecule of the previously existing gas is diminished by compression, so the inflow of new purchasing power into the economic system will compress the old purchasing power. When the price changes which thus become necessary are completed, any given commodities exchange for the new units of purchasing power on the same terms as for the old, only the units of purchasing power now existing are all smaller than those existing before and their distribution among individuals has been shifted. This may be called credit inflation. (Schumpeter 1934, p. 109)

However, this credit inflation is normally only temporary:

After completing his business - in our conception, therefore, after a period at the end of which his products are on the market and his productive goods used up - he has, if everything has gone according to expectations, enriched the social stream with goods whose total price is greater than the credit received and than the total price of the goods directly and indirectly used up by him. Hence the equivalence between the money and commodity streams is more than restored, the credit inflation more than eliminated, the effect upon prices more than compensated for, so that it may be said that there is no credit inflation at all in this case - rather deflation - but only a non-synchronous appearance of purchasing power and of the commodities corresponding to it, which temporarily produces the semblance of inflation. Furthermore, the entrepreneur can now repay his debt (amount credited plus interest) at his bank and normally still retain a credit balance (=entrepreneurial profit) that is withdrawn from the purchasing-power fund of the circular flow. Only this profit and interest necessarily remain in circulation; the original bank credit has disappeared ... (op. cit., pp. 110–11)

Schumpeter argues that credit inflation is thus temporary if the investment projects are successful. In fact, innovation should lead to a secular fall in the price level as less costly production methods are employed.

For the individual bank, the limit to its ability to create credit is set by the constraint that credit inflation must be temporary. In other words, the credit it grants must be (eventually) retired. Credit inflation, if it were permanent, would ruin a bank (that is, if loans are never repaid). However, if all banks together continually refinanced credit (making bad loans good by extending debt repayment schedules), permanent credit inflation would result, which would not ruin the banks. In this case, aggregate price levels would continually rise, leading to trade deficits and gold outflows. Schumpeter thus argues that some legal restrictions might be required to prevent banks from operating to generate permanent credit inflation. (op. cit., pp. 114–15)

He argues that credit is unnecessary to maintain a circular flow and is, therefore, unnecessary in an economy that is not growing. Since he relates growth to entrepreneurial innovation, the essential function of credit is to allow development through innovation. He contends that the money market develops with

the use of credit to finance innovation, but the money market is not essential merely to circulate a given flow.

And so the requirements of the circular flow are added to the entrepreneur's demand in the money market on the one hand, and money from the circular flow increases the supply in the money market on the other. Hence we feel in every money-market article the pulse of the circular flow ... But this must not prevent us from distinguishing the transactions in the money market which belong to the circular flow from others. Only the latter are fundamental; the former are added onto them, and the fact that they appear in the money market at all is merely a consequence of development. (Schumpeter 1934, pp. 124–5)

Financial institutions are also entrepreneurial organizations that strive to innovate to generate capital gains.

Financial institutions and practices enter our circle of problems in three ways: they are “auxiliary and conditioning”; banking may be the object of entrepreneurial activity, that is to say, the introduction of new banking practices may constitute enterprise; and bankers (or other “financiers”) may use the means at their command in order to embark upon commercial and industrial enterprises themselves. (Schumpeter 1951, p. 222)

A bank is a business that makes profits by lending its promises to pay. Schumpeter speculates that three-quarters of a bank's deposits are merely credits, and “there can be no doubt that these circulating media come into being in the process of granting credit and are created especially ... for the purpose of granting credit.” (Schumpeter 1934, p. 98) He argues that entrepreneurs borrow what they deposit: bank credits. Like other entrepreneurs, those in the banking industry earn capital gains by innovating, that is, by finding new and less costly ways of granting credit. In credit markets, the entrepreneurs are the demanders and banks are the suppliers.

The creation of credit is merely “the exchange of present against future purchasing power” (the borrower obtains purchasing power today by promising to surrender purchasing power in the future). (op. cit., p. 125) The main purpose of credit markets is to finance development. The financial institution earns profits from the interest payments made by entrepreneurs who have borrowed and successfully undertaken a project. The financial entrepreneur earns capital gains by finding innovative financing methods that provide the purchasing power needed by other entrepreneurs.

In summary, credit plays a central role in a capitalist society because it funds growth and innovation. Banks and other financial institutions provide this credit. They are profit-seeking organizations, whose only quantity constraints are determined by the necessity that they earn profits. The most important determinant of profits is the necessity that “credit inflation” be temporary: that

is, that loans be repaid. Financial institutions that engage in “entrepreneurial activity” earn above-normal profits (capital gains) by innovating.

Schumpeter’s analysis is consistent with the endogenous money approach. Banks create credit on demand, and this credit enables expenditure to proceed. The growth of net investment requires the creation of credit (growth cannot be financed out of prior savings). Banks have an incentive to meet the demands for credit to earn profits. Banks that innovate can reduce costs and earn higher profits. Thus, money is created during the normal capitalist processes (including investment) by profit-seeking financial institutions (lenders) to meet the demand of profit-seeking nonfinancial institutions (borrowers).

Schumpeter distinguishes between capitalists and entrepreneurs. Capitalists own productive capacity that generates revenue they can draw on to finance their production process. A given circular flow of credit can finance their activities. Entrepreneurs are innovators who are not yet producing (at least in their new innovative line of business); thus, new finance must be created—beyond the existing circular flow. This does not mean that capitalists do not need credit—it simply means that they do not need a growth of credit. Because his focus is on the innovation process, he has simplified what is going on in the existing circular flow. Over any short period, there will be capitalists whose revenue does not match spending, necessitating bridge finance. Further, capitalism is continually growing even without innovations—simply expanding the scale of employment and output of existing lines of production. Also, interest rates can rise, necessitating greater spending on interest costs. Prices, wages, costs, and profit rates can all rise, also necessitating the expansion of credit. The simple circular flow—the subject of the Circuit Approach examined in the next chapter—can be expanded to deal with these topics. But for Schumpeter’s purpose, it was not necessary to do so.

NOTES

1. See Panico 1982, p. 20.
2. The term money as used here is synonymous with government fiat money.
3. On the other hand, Thornton provided a summation of the proper role for the central bank which would have been accepted by the Currency School: “To limit the total amount of paper issued, and to resort for this purpose whenever the temptation to borrow is strong, to some effectual principle of restriction ... to afford a slow and cautious extension of it, as the general trade of the kingdom enlarges itself ... and to lean to the side of diminution, in the case of gold going abroad, and of the general exchanges continuing long unfavourable; this seems to be the true policy of the directors of an institution circumstanced like that of the Bank of England.” (Henry Thornton, quoted in Mints 1945, p. 56)

4. Throughout this article, Keynes emphasized the role of credit in financing an increase in investment. (Keynes 1973, pp. 278–85)
5. In his defense of the *General Theory*, Keynes argued that he had not adopted a fixed money supply. See Keynes 1973, p. 232.
6. Note that this differs from Tooke's theory, in which credit is granted to circulate existing goods, so it cannot be inflationary.

6. Endogenous money: a credit and state money approach

INTRODUCTION

After Keynes, an exogenous money supply was adopted in mainstream theory and was incorporated within IS-LM analysis as well as Friedman's monetarism. Over the years, this position has been challenged by heterodox economists at various times until it was finally abandoned by most mainstream economists in the 1990s. This chapter summarizes several of these challenges, beginning with that of the Radcliffe Committee, to draw out the endogenous money thread that links the works of the economists analyzed here. The final section integrates endogenous credit money and a Chartalist approach to the government's money. While some critics of MMT have argued that the state money approach is not consistent with a credit money approach, we argue that both the history of money and the evolution of heterodox thinking about money require a credit *and* state money approach.

KALDOR AND THE RADCLIFFE COMMITTEE

In 1959, the Radcliffe Committee in Britain issued its report: *The Working of the Monetary System*. It analyzed the institutional changes over the previous decades to determine the impact these had on conventional monetary policy. Specifically, the committee set out to determine whether tight monetary policy and high interest rates could be used to contain high aggregate demand. It was feared that high levels of liquidity after the war (due to forced saving) would lead to excessive spending. Thus, tight monetary policy might be used to forestall inflationary pressures.

The Radcliffe Committee, however, concluded that tight monetary policy would not prevent spending but would probably cause distortions of spending due to high interest rates that would slow investment and reduce technological innovation. The Committee argued:

Spending is not limited by the amount of money in existence; ... it is related to the amount of money people think they can get hold of, whether by receipts of income (for instance from sales), by disposal of capital assets or by borrowing ... spending

is not limited by the amount of money in existence ... [due to variability of] the velocity of circulation. (Radcliffe Committee 1959, p. 133)

The Committee concluded, "We cannot find any reason for supposing, or any experience in monetary history indicating, that there is any limit to the velocity of circulation of money." (ibid.) Furthermore, "the more efficient the financial structure, the more can the velocity of circulation be stretched without serious inconvenience being caused." (ibid.)

The emphasis of the Committee on variable velocity may have been due to Kaldor's testimony. He presented evidence showing that velocity increased between 1955 and 1958 and attributed it directly to tight money policy that had not allowed the money supply to expand as rapidly as spending. He argued that there is no direct impact of money on income since velocity can expand without limit. Instead, as the report stated, "the impact effect of any change in the money supply is not on the level of payments at all, but on the velocity of circulation." (Radcliffe Committee 1959, p. 146) Thus, tight money would be an ineffective means of controlling inflation, working only if interest rates were pushed high enough to induce a deep recession. He argued that a smaller change in interest rates might have no significant impact on spending since the investment-interest rate link was not close. Furthermore, while a large change in interest rates might lower investment, that would tend to shift spending toward the consumption share, with adverse consequences for long-term growth.

Thus, Kaldor and the Radcliffe Committee primarily relied on variable velocity as a critique of the quantity theory of money. Tight monetary policy could indirectly affect spending only if interest rates were raised sufficiently. However, by 1970 Kaldor realized that the emphasis on velocity had been misplaced. He began to argue that the money supply itself (and not just velocity) is endogenously determined. (Kaldor 1970) Rather than arguing that velocity would adapt to a rise in expenditures, he argued, "the increase in the supply of money is a consequence of increased loan expenditure." (Kaldor and Trevithick 1981, p. 5) When firms utilize credit lines, "there will be an automatic increase in the money supply for the simple reason that additional expenditure will swell the bank deposits of the recipient." (op. cit., p. 6)

He thus argued that the demand for loans determines the supply of money because bank lending is limited only by the credit demands of worthy borrowers. If a stable relation exists between expenditures and the money supply (that is, a constant velocity), this is simply proof that expenditures raise the level of the money supply when they are based on the extension of credit. (Kaldor 1985, p. 22) Central bank control over the money supply consists of its ability to force interest rates to rise, which might reduce expenditures and income.

The fall in the money supply is then only a result of the fall of expenditures. (op. cit., pp. 22–5)

Furthermore, he argued that a tight money policy might even increase the money supply as well. This is because higher interest rates would increase the attractiveness of holding various interest-bearing deposits relative to other assets. (Kaldor 1985, p. 75) The attractiveness of bank deposits enables banks to make loans by creating deposits without worrying about a clearing drain as these deposits are exchanged for nonbank liabilities. As evidence, he pointed to the Monetarist experiments in the United Kingdom and the United States that began in 1979, in which an extremely tight money policy raised interest rates but increased the rate of growth of the money supply. (op. cit., pp. xv, 75, 97)

Monetarists (like Friedman) countered that the failure of Chairman Volcker's "Monetarist experiment" should be attributed to the incompetence of the central bank, and not to Monetarist theory. If only the monetary authorities had been competent, they could have hit money targets and eliminated inflation without inducing a recession. Kaldor responded that these targets cannot be hit because tight money induces innovation and economization of cash, thereby increasing the ability of banks to attract deposits (at the expense of other financial assets) at the higher interest rates. (Kaldor 1985, pp. xiii, 46) Finally, the central bank must ultimately act as a lender of last resort so that it cannot close its discount window merely because the rate of growth of the money supply exceeds targets without endangering the entire banking system. (op. cit., p. 25) However, the central bank is free to establish the discount rate and thereby affect the entire term structure of interest rates: "The rate of interest ... is the one instrument which is entirely under the control of the Government." (op. cit., p. 97)

Kaldor thus offers a much stronger criticism of Monetarism than does the original Radcliffe thesis. Friedman believed that the Radcliffe position (and, thus, Keynesianism) relied on the hypothesis of an unstable velocity of money. Thus, when he was able to show stability of velocity through the 1960s, he believed that this clearly contradicted the Keynesian position and lent support to the quantity theory of money. However, with an endogenously determined money supply, stability of velocity is also perfectly consistent with Keynesian theory. In Kaldor's view, spending determines money, but there is also some feedback from monetary policy to spending through the interest rate.

Kaldor's position is most consistent with that of Moore's, to be examined in detail below. Both Kaldor and Moore argue that banks are never quantity constrained because the central bank ultimately accommodates the bank's demand for reserves. Like Moore, Kaldor argues that the central bank determines the short-term interest rate (fed funds rate in the US), which is set by a markup over the central bank's discount rate. Kaldor argues:

One cannot conceive of the short rate being “determined” in any other way than through the rediscount rate or the open market policy of the Central Bank. Indeed, it is only through their power to control the whole range of short-term interest rates that the monetary authorities can be said to “control” the supply of money in its broader sense. (Kaldor 1980, p. 6)

Like Moore, Kaldor relies critically on accommodative central bank policy. Indeed, he argues that before World War I, it was easy for central banks to control the money supply, whose quantity was exogenously determined by gold reserves. (Kaldor 1985, p. 102) Endogenous money has existed only since WWI, after which central banks began to accommodate the demand for reserves rather than forcing banks to adapt to varying gold stocks. However, this is not consistent with the arguments presented in Chapters 2 and 3 above—banks and other issuers of liquid liabilities have always been able to get around constraints.

HYMAN MINSKY AND THE FINANCIAL INSTABILITY HYPOTHESIS

Minsky’s earliest work in the 1950s can also be seen as consistent with an endogenous money approach, and, indeed, his later work on financial instability requires it. Minsky’s original dissertation advisor was Schumpeter, who heavily influenced Minsky’s work on money and finance. He extended Schumpeter’s theory of the importance of innovation to the financial sector itself. In 1957, Minsky related institutional innovation to profit opportunities and showed how innovation by banks allows business activity to expand even without an expansion of bank reserves. Specifically, he showed how the development of the fed funds market allows a given level of aggregate reserves to support a greater expansion of deposits, and how repurchase agreements allow a given quantity of demand deposits to support a greater volume of loans. (Minsky 1957)

Minsky also related innovation to tight monetary policy. If a business expansion causes the monetary authority to fear inflation, it institutes tight monetary policy, which leads to a rise in interest rates. Tightness in the money market doesn’t, however, immediately reduce the demand for loans if expectations are optimistic. Furthermore, as interest rates rise, liquid balances are reduced to take advantage of profit opportunities so that the velocity of liquid assets increases. Financial institutions also want to take advantage of the higher rates and so find new ways of meeting loan demand. Innovations result in a shift in the relation between velocity and interest rates: a higher velocity will be associated with each interest rate. Thus, central bank constraints raise interest rates, inducing innovations that increase velocity. Monetary policy won’t be

effective in constraining monetary aggregates unless it decreases aggregate reserves sufficiently to compensate for the rise in velocity. (Minsky 1957)

Innovations are related to Minsky's financial instability hypothesis. When innovations are induced by a demand for finance that can't be met within the existing institutional constraints, systemic liquidity is reduced. "(E)very institutional innovation which results in both new ways to finance business and new substitutes for cash assets decreases the liquidity of the economy." (Minsky 1957, p. 184) Innovations allow leveraging of liquidity as increasingly illiquid assets are added relative to cash, government bonds, or short-term bank debt in portfolios. This increases the potential for instability of the system since failure by only a few large firms (financial or non-financial) to meet debt commitments can lead to a cascade of failures as assets are written down.

Early in a boom, high demand for finance will not raise interest rates much since banks and the money market have elastic supply curves, and firms and households will take illiquid positions to receive a higher return if they are optimistic about the future. But declining liquidity (rising debt to net worth, or to safe asset, ratios) increases instability by exposing assets to the possibility of rapid deflation of value should some units fail to meet commitments. Ultimately, the major limiting factor regarding how far the value of financial assets can fall is the willingness of the central bank to purchase them. Thus, the monetary authorities can halt a debt deflation process by intervening as a lender of last resort to increase the liquidity of the system by accepting or lending against illiquid assets. (Minsky 1957, p. 185)

The central bank cannot exogenously set the money supply. In an expansion, producers voluntarily reduce liquidity positions, allowing production to expand even if the money supply (narrowly defined) does not expand. However, it is likely that the money supply (even narrowly defined) will expand as innovations permit the expansion of loans and demand deposits on a given reserve base. Furthermore, as is now universally recognized, central banks operate with an interest rate target (even though they announced money growth targets in the Monetarist experiment, they now clearly announce interest rate targets). That means they must supply reserves to keep the overnight rate (fed funds in the US) from rising above the target.

In a crisis, the desired liquid position may not be reached until the value of assets has fallen so far that employment and income levels have dropped catastrophically. This is the Fisher debt deflation case, something central banks do not want to happen. To prevent depression, the central bank can supply liquidity (for example, by increasing the reserve base) so that liquid positions can be restored before asset values and employment collapse. The central bank sets a floor to the value of assets by acting as a lender of last resort and/or by directly intervening to guarantee the value of bank assets (for example, through direct

purchases of assets). Thus, the central bank must ultimately operate as a lender of last resort to prevent debt deflation.

After 1957, Minsky gradually developed a sophisticated view of the relations among credit, asset prices, and instability. According to Minsky, rising expectations raise the demand price of capital assets above their supply price, inducing investment. Some of the investment is financed out of internal profit flows, but much is financed externally by financial institutions. As an expansion proceeds, fueled by investment, profits rise and reward units that have expanded their balance sheets.¹ Firms become more willing to commit a larger portion of expected income flows to servicing debt, while banks become more willing to leverage equity as they provide loans and create deposits. Financial fragility increases as income flows and payment commitments become increasingly synchronous. A crisis can occur if finance costs rise, if liquidity preference rises, or if income flows turn out to be less than expected. Endogenous processes tend to ensure that one of these (or all three) will, in fact, occur. (Minsky 1978, 1986)

Minsky argues that as bank balance sheets expand, there is upward pressure on interest rates because banks are concerned with leverage ratios. Furthermore, the central bank tends to raise interest rates near the peak of an expansion. (Fazzari and Minsky 1984) Rising interest rates increase the portion of income flows that firms must commit to debt payments (for example, firms must roll over some of the debt at higher rates since construction projects take time). As interest rates rise, “present value reversals” (projects that were expected to be profitable at lower interest rates become unprofitable at higher rate) occur, which reduces planned investment (and may even cause some projects to be abandoned). (Minsky 1978, p. 16, and 1986, p. 195) As investment declines, aggregate profit flows decline and make it even more difficult to meet payment commitments.

Minsky does not present a strict dichotomy between banks and financial institutions, nor between financial institutions and nonfinancial institutions. All types of economic institutions engage in balance sheet operations, and any economic entity can create money in the sense of issuing a money-denominated liability. As Minsky says, the problem lies in getting it accepted. (Minsky 1986, p. 228) Liabilities are used to finance positions in assets and represent a “money now for money later” proposition. A wide variety of balance sheet items enable one to spend now and pay later. (Minsky 1986, p. 228)

Certain liabilities of financial institutions also circulate as media of exchange and means of payment. However, according to Minsky, banks are not intermediaries, but enhance the debts of customers. “Banking is not money lending; to lend, a money lender must have money. The fundamental banking activity is accepting, that is, guaranteeing that some party is creditworthy ...” (ibid., p. 229) Minsky used to joke that “money lenders” stand on the street

corner in Chicago wearing trench coats with pockets full of cash. They lend the cash at high interest rates, and the penalty for failing to make a timely payment is severe physical punishment. The point is that banks create what they lend—their own deposits. They do not collect money from depositors that they can lend out.²

Minsky's insights add significantly to Kaldor's argument. In an expansion, bankers share optimistic profit expectations with firms. Firms and banks are willing to increase leverage ratios, so the ratio of debts to liquid assets rises as firms borrow to finance investment and banks extend finance. Even if the central bank tries to limit the expansion of some monetary aggregate (for example, reserves or M1), this has little effect in an expansion since firms and banks willingly stretch liquidity. In the early stages of an expansion, balance sheets are expanded even with no rise in interest rates because innovations create new ways to finance positions in assets and, with robust expectations, firms are willing to increase leverage. When institutional constraints are reached, however, further expansion might require rising interest rates. Rising interest rates, in turn, increase systemic fragility because a greater portion of income flows must be committed to meeting finance payments. If interest rates rise so high that a few large firms decide to decrease spending, or if a few large firms are unable to meet commitments, the economy slows and expectations fall. Economic entities revise desired leverage ratios downward. However, general liquidation of positions in assets is not possible without declining asset values precisely because the expansion was based on leveraged purchases of assets. Nevertheless, government deficits place a floor on the level of profit flows, and central bank intervention as a lender of last resort places a floor to the value of bank assets and liabilities. (Minsky 1978, p. 17)

One element of Minsky's argument now seems outdated: he did argue that central banks had some discretion over the amount of reserves supplied. When Minsky developed his arguments, the Fed and other central banks did not announce interest rate targets. During the early 1980s Monetarist experiment, they claimed to be targeting the money supply. Since 1994, however, the Fed has explicitly adopted and announced its interest rate target. With an interest rate target, the Fed cannot leave banks short of reserves—they will bid the fed funds rate above target and the Fed must provide the desired reserves to make sure it doesn't miss its own target. Alternatively, if the Fed leaves too many reserves in the banking system, the overnight rate will decline—potentially to zero. However, since 2009 the Fed pays interest on reserves, so the rate it pays sets the lower bound. This allows the Fed to put excess reserves into the system that earn the rate paid by the Fed. In sum, the Fed cannot leave banks short—it must accommodate the demand—but can provide more reserves than banks want. What all of this means for Minsky's approach is that we should focus

on interest rate policy (i.e., the price of reserves) rather than on the quantity of reserves. This is the approach of the Horizontalists, examined below.

JAMES TOBIN, THE NEW VIEW, AND THE MONETARIST COUNTERATTACK

In 1963, in “Commercial Banks as Creators of Money,” Tobin distinguished between the “old view” and the “new view” of banking. (Tobin 1987) According to the old view, there is a clear distinction between banks and other types of financial institutions, and between money and other types of financial assets. Only commercial banks can create money, and they do so through the deposit expansion process. Since they possess “the widow’s cruse,” they “have to be restrained by reserve requirements. Once this is done, determination of the aggregate volume of bank deposits is just a matter of accounting and arithmetic: simply divide the available supply of bank reserves by the required reserve ratio.” (Tobin 1987, p. 273)

The new view, on the other hand, focuses “on demands for and supplies of the whole spectrum of assets rather than on the quantity and velocity of ‘money’ ...” (op. cit., p. 274) In addition, the focus is on the full range of financial institutions.

(T)he essential function of financial intermediaries, including commercial banks, is to satisfy simultaneously the portfolio preferences of two types of individuals or firms. On one side are borrowers, who wish to expand their holdings of real assets - inventories, residential real estate, productive plant and equipment, etc - beyond the limits of their own net worth. On the other side are lenders, who wish to hold part or all of their net worth in assets of stable money value with negligible risk of default. (ibid.)

Intermediaries allow borrowers to obtain funds to purchase real assets at a lower interest rate than if they were to borrow directly from savers. Any increase in “the amount of financial intermediation in the economy can be expected to be, *ceteris paribus*, an expansionary influence.” (op. cit., p. 275) Commercial banks are special since their liabilities function as means of payment. However, since there is a high degree of substitutability among the liabilities of financial intermediaries, other financial intermediaries also influence real spending.

All financial intermediaries are limited in their ability to expand balance sheets. “Neither individually nor collectively do commercial banks possess a widow’s cruse.” (op. cit., p. 276) Expansion can occur only up to the point where “the marginal returns on lending and investing, account taken of the risks and administrative costs involved, will not exceed the marginal cost to the banks of attracting and holding additional deposits.” (op. cit., p. 277) In

the absence of reserve requirements, “expansion of credit and deposits by the commercial banking system would be limited by the availability of assets at yields sufficient to compensate banks for the costs of attracting and holding the corresponding deposits.” (op. cit., p. 279) But in the presence of required reserve ratios (and particularly if deposit interest rates are controlled by legal ceilings), an expansion of reserves will normally lead to an expansion of loans since “the marginal yield of bank loans and investments exceeds the marginal cost of deposits to the banking system.” (ibid.)

The simple deposit multiplier doesn’t necessarily apply to all expansions of reserves. Banks may decide to hold excess reserves if the interest rates on various types of assets aren’t sufficient to cover the costs of obtaining funds. A rise in bank reserves may induce a rise in the quantity of loans extended, but this will also tend to reduce the interest rate that can be charged, which will “diminish the incentives of banks to keep fully loaned up or to borrow reserves, and to make banks content to hold the average higher excess reserves.” (op. cit., p. 280)

Much of what Tobin has argued is consistent with the endogenous money approach. However, there are several points of conflict. First, Tobin argues that he does not mean to imply “that actual monetary authorities have any less control than textbook monetary authorities.” (op. cit., p. 289) While there is some “slippage” between reserves and liability expansion, and between the quantity of financial assets and real expenditures, the monetary authority can offset these slippages through manipulation of reserves. Thus, by controlling the base, the central bank can control the money supply. Second, although Tobin realizes that changes in depositor preferences can affect the ability of banks to make loans, he does not appear to fully recognize the extent to which bank liability management frees banks from quantity constraints. And, finally, Tobin does not seem to acknowledge that any extension of credit necessarily creates the deposits that the system as a whole requires—any bank losing deposits can obtain them in the “wholesale” markets. Asset and liability management, innovations, and central bank accommodative behavior essentially ensure that banks can lend if creditworthy borrowers exist.

Tobin’s 1963 paper led to other work that further developed the “new view,” as an alternative to the Monetarist view that was rapidly gaining adherents during the 1960s. Gramley and Chase (1965), Karekan (1967), Cacy (1968), and Davis (1968) extended Tobin’s analysis, including the role played by both real and financial assets, as well as by interest rate movements. They also stressed that the money supply is a function of a complex series of relations among banks, the public, and the central bank. Thus, they dismissed the simplistic view that the money supply is determined by the base (set by the central bank) multiplied by a stable money multiplier.

Brunner (1968) attacked the new view on several levels. First, he argued that the new view's emphasis on "a systematic application of economic analysis, in particular an application of relative price theory, to the array of financial intermediaries, their assets and liabilities" (Brunner 1968, p. 11) did not distinguish it from the Monetarist approach, in which "monetary impulses are transmitted to the economy by a relative price process which operates on money, financial assets (and liabilities), real assets, yields on assets and the production of new assets, liabilities, and consumables." (op. cit., p. 18) Second, the new view made "little attempt at specific structuring and empirical content." (op. cit., p. 12) The new view abstracts from the observable world to such a great extent that it is nearly empty of form. In fact, Brunner argued that much of the new view cannot be empirically tested, but empirical evidence contradicts those hypotheses that are testable.

Brunner argued that the new view can be reduced to the hypothesis that the money multiplier is variable (note the similarity to Kaldor's argument that velocity of money is variable) and that it could easily be tested by a regression of the money stock on the base. If the money multiplier is stable, then the new view would not be supported. Brunner found that over 90 percent of the variation in the money stock is explained by changes in the monetary base. Furthermore, he argued that changes in the base can be directly controlled by the central bank. "In summary, preliminary investigations yield no support for the contention that the behavior of banks and the public dominates cyclical movements in the money stock." (op. cit., p. 18) He admitted that "the existence of a mutual interaction over the shorter run between money and economic activity ... must be fully acknowledged. Yet, this interaction results from the conception guiding policymakers which induces them to accelerate the monetary base whenever pressures on interest rates mount ..." (op. cit., p. 20) In other words, if the central bank did not target interest rates, there would be little feedback from economic activity to the money stock—with the primary causal sequence going from central bank control of reserves, to control of the money supply, and hence to control of economic activity. Thus, Brunner was able to incorporate some of the new view into Monetarism, while rejecting the rest.

Ultimately, the new view was not an effective critique of Monetarism because it did not develop an adequate theory of the way in which money enters the economy. The new view's emphasis on portfolio adjustments due to changing interest rates was easily incorporated within Monetarism. Because the new view accepted that the base is a determinant of the money supply, Monetarists were able to reduce the new view to the hypothesis that the money multiplier is variable. When Monetarists were able to show a high correlation between the money stock and the base, they believed they had effectively disproven the new view (or at least, had relegated it to a theory of minor fluctuations in

the money multiplier). The results were similar to the Monetarist critique of Kaldor's "Keynesian" theory: they believed that if they could show that the income velocity of money is fairly stable, then the quantity theory of money would be supported, disproving Keynesian theory. (Kaldor 1985)

However, neither a stable money multiplier nor a stable income velocity of money is inconsistent with the endogenous approach to money. If money is created to finance spending, and if the expansion of bank balance sheets requires an expansion of reserves, then stable relations between spending and money, and between money and reserves, will exist. Thus, Brunner's empirical tests cannot distinguish between the endogenous approach and Monetarism. Tobin and the other new view economists did not go far enough in their development of an alternative to Monetarism because they accepted a direction of causation running from the base to money, and because they believed the central bank can control the base—at least in theory.

As we will see below, Moore's Horizontalist approach can deal with both of the Monetarist responses. But first, we will examine one other earlier attempt to tackle the topic of money and finance.

GURLEY AND SHAW: INTERMEDIATION AND DIVERSIFICATION

Gurley and Shaw (1955, 1956, 1960) began with an analysis of the finance process. Spending may be financed internally (through income flows) or externally (through debt). External finance includes direct and indirect finance. Direct external finance occurs when a surplus unit (one whose expenditures are less than its income) acquires financial claims on debtors. Indirect external finance occurs when surplus units purchase claims on financial intermediaries, which then make loans to the ultimate borrowers. (Gurley and Shaw 1956, pp. 258–9) In a closed economy with no government, aggregate income equals expenditure, so that net borrowing must be zero. Each private unit that spends more than its income (incurring debt) is offset by a surplus unit (that acquires financial assets). Net financial private wealth must be zero in a closed, private economy. (Gurley and Shaw 1960) (This is related to the sectoral balance approach of Wynne Godley, which is essential to the arguments made by MMT about government finance. We will return to that in Chapter 8.)

However, there are two possible sources of private sector net financial wealth: a government sector and a foreign sector. One country can have positive net financial worth if it obtains net credits on another. Furthermore, the government sector can issue liabilities that become a source of net wealth for the private sector. Gurley and Shaw then distinguished between inside and outside money. Inside money is the creation of the private sector and is offset by a private liability (note that the government could hold financial liabilities

of the private sector, and these would need to be subtracted from the government's financial liabilities held by the private sector to calculate private sector net financial wealth). Thus, when a bank makes a loan, it creates money that is an asset of the nonbank sector, but also creates a liability of the banking sector, and the nonbank sector issues a liability (the IOU) that is held as the asset of the banking sector. Therefore, inside money cannot be a net source of wealth to the private sector. However, government debt (bonds or money) is a net source of wealth for the private sector. The money issued by the government is outside money. (Gurley and Shaw 1960, pp. 72–5)

Gurley and Shaw noted that orthodox theory (especially Monetarism) draws a sharp distinction between the “monetary system” (commercial banks and the Fed) and other financial institutions. In orthodox theory, the monetary system issues money while financial “intermediaries” merely “transfer to investors any part of this money supply that may be deposited with them by savers.” (Gurley and Shaw 1956, p. 260) However, Gurley and Shaw argued that “the monetary system is in some significant degree competitive with other financial intermediaries.” (op. cit., p. 261) Financial intermediaries issue debt that competes with money issued by the monetary system: “their issues of indirect debt displace money ...” (ibid.) The debt of financial intermediaries functions like money until it becomes necessary to make monetary payments. That is, “other financial assets can be accumulated preparatory to money payments, as a precaution against contingencies, or as an alternative to primary securities.” (ibid., p. 261)

Since the liabilities of a wide variety of financial institutions act as substitutes, “for any level of income, the money supply is indeterminate until one knows the degree of substitutability between money created by banks and financial assets created by other intermediaries.” (op. cit., p. 261) When spending by deficit units rises, financial assets and liabilities are created. To the extent that surplus units are willing to accept “near monies” (liabilities of nonbanks) rather than money (liabilities of the “monetary system”), spending can rise with no increase in the money supply (narrowly defined). That is, the velocity of (narrow) money rises. Gurley and Shaw explain:

(I)t is a mistake to suppose that the optimal growth of the money is some simple function of trends in income or even of income and debt. The necessary growth in the money supply may be high or low, positive or even negative, depending on the growth of income, the share of spending that is externally financed (especially by long-term securities), the growth in demand by spending units for direct relative to indirect financial assets, and on the development of financial intermediaries whose indirect debt issues are competitive with money. (Gurley and Shaw 1955, p. 531)

If rising spending by deficit units were financed solely by direct finance, surplus units would likely require rising interest rates to induce them to hold

increasing quantities of direct debt. This would occur, for example, if there were a preference for liquidity since the money supply would be constant, while illiquid forms of assets and liabilities would be rising. Thus, continued expansion of spending might require that financial intermediaries intervene, purchasing the direct debts of deficit units and issuing indirect debt to surplus units. Eventually, surplus units might acquire all the indirect debt of financial intermediaries desired. In this case, interest rates on these liabilities will have to rise. Alternatively, the monetary system can intervene, with commercial banks acquiring direct debt and issuing money to satisfy the liquidity preferences of surplus units.

Finally, diversified portfolios are desirable to reduce risk.

In brief, the demand for money, the demand for diversification among other financial assets, offers of differentiated primary securities, and the demand for mixed asset-debt positions are to be explained by a common principle. They are tactics of risk-avoidance by spending units in a society where division of labor between saving and investment creates stocks of claims and counterclaims. (Gurley and Shaw 1960, p. 121)

Because the government creates outside money, policy is not neutral. A rise in outside money raises money balances, and since this is not offset by a rise in private debt, there is an increase in the financial net wealth of the private sector. Furthermore, outside money increases relative to other financial assets. As surplus units attempt to diversify, the interest rate can be driven down on direct and indirect liabilities of deficit units. This, in turn, stimulates expenditure and production. In the absence of government debt, the private sector might reach a point at which further expansion is inhibited because the liquidity premium required to induce surplus units to hold debt is too great. But, an increase in outside money can alter the premium since liquidity rises.

Inside money is not neutral, either. Gurley and Shaw call their approach the gross-money doctrine, in which both inside and outside money matter. They counterpose it with the net-money doctrine, in which only outside money affects real variables.

[N]et money doctrine implies that the quantity of inside debt and its counterparts in financial assets have no net effects on the aggregate real demand for money, but this overlooks the desire for diversified financial positions by both firms and consumers ... The doctrine distinguishes too sharply between the behavior of private domestic sectors and that of outside sectors. Stockpiling of financial assets is a defensive measure for spending units in a hostile economy, much as stockpiling of weapons is a defensive measure for nations in a hostile world. And in each case diversification of the stock usually pays off. Spending units' demand for money, as one financial asset among many, depends partly on the types of frictions and uncertainties associated with alternative financial assets. (Gurley and Shaw 1960, p. 188)

Gurley and Shaw's distinction between inside and outside money has been attacked on two fronts. First, it has been argued that government debt is not outside money. According to Gurley and Shaw, when the government issues bonds to the central bank to finance spending, the quantity of outside money rises because the central bank issues liabilities that are offset by government debt rather than private debt. However, the critique argues that the public should realize that this government debt represents a future claim on private income (in the form of future taxes—the so-called Ricardian Equivalence doctrine). Thus, there is no net increase in private wealth. However, the rate of discount applied to future taxes might differ from the rate of discount applied to future interest receipts so that the present value of government debt may be more than the present value of future tax payments. Finally, it is doubtful that the public believes that interest will be paid out of future tax revenues. This disbelief is not irrational if one can use the past as a guide to the future (in the USA, the debt-to-GDP ratio has been growing on trend since the nation's founding). In any event, as we will argue in Chapter 8, government and its debt really *is* different—it does not represent a claim on future generations.

The second critique of Gurley and Shaw takes the opposite view: all money is outside money. Money is characterized as the capitalized value of the returns it yields in services. When banks make loans and create demand deposits, these deposits function as money and provide utility to their holders. This real increment to utility is not offset by a decline in the bank's utility since it doesn't give up any money. Harris summarizes this position:

In Pesek and Saving's analysis, therefore, the creation of demand deposits by the bank giving loans to the private sector does create net wealth. The private sector gains an asset (money) while the banking sector does not increase its liabilities except in a bookkeeping sense, since demand deposits on which no interest is paid are not a genuine liability to the banking sector. (Harris 1981, p. 39)

This argument ignores uncertainty. By accepting the loan, the bank has increased the ratio of risky assets to equity and liquid reserves. If the loan turns out to have been to a creditworthy customer, then equity is increased (as the interest earned by the bank is greater than the interest it pays out); otherwise, the bank suffers a fall in equity. Since the future is unknowable, it is impossible for the bank to eliminate (or compensate for) uncertainty entirely. Furthermore, by creating demand deposits, the bank has incurred a monetary liability. Should the depositor decide to hold cash, or should the deposits flow to another bank, the bank is required to come up with cash or reserves. The public's gain in liquidity is offset (at least in part) by the bank's loss. Since utility is immeasurable and cannot be compared across economic units, it is impossible to state whether there is a net increase in utility. However, the rise

in “utility-yielding services” that results when banks create deposits is accompanied by a rise in bank “disutility” since banks are ultimately responsible for providing liquidity on demand. They will either need to borrow from other banks at the market rate or from the central bank’s discount window (facing potential frown costs).

In any case, the Pesek and Saving (1967) argument is irrelevant to the major conclusion of Gurley and Shaw, which is that both inside and outside money matter. The purpose of Gurley and Shaw’s thesis was not to show that only outside money allows real output to increase. Indeed, they consistently argued that there is a high degree of substitutability among various financial assets, so that real spending can rise even if the money supply (whether defined narrowly as government debt in the form of outside money, more broadly as government debt plus commercial bank deposits, or even more broadly to include deposits at other financial institutions) is held constant. Credit expands to allow some units to spend more than income. Of course, expenditure creates income, so the deficit of the spender is offset by a surplus. Where surplus units do not wish to hold all the direct debt issues, financial intermediaries issue indirect debt, some of which may include money (however defined). As production increases, the demand for money may rise, so that continued expansion with constant interest rates eventually requires expansion of the money supply. However, the linkage between money and income is very imprecise. As Gurley and Shaw say, “The same money supply and national income are compatible with various interest rate levels, depending upon the size of non-monetary intermediaries and upon the degree to which their issues are competitive with money.” (Gurley and Shaw 1956, pp. 263–4)

Gurley and Shaw’s analysis of financial institutions is consistent with the endogenous money approach. Spending is not limited by a narrowly defined monetary aggregate since financial intermediaries offer various liabilities that allow spending to rise. Gurley and Shaw’s discussion of portfolio diversification adds to the endogenous money approach. If firms had to issue securities directly to surplus units, then an increase in spending might require rising interest rates because economic units prefer to diversify portfolios. However, financial intermediaries provide liquid liabilities so that surplus units *can* diversify. This allows investment to proceed without causing interest rates to rise. In this way, a given quantity of liquid assets (including inside and outside money) can support a wide range of spending levels and inside money.

MOORE: THE HORIZONTALIST APPROACH

Moore argues that only when money is metallic money is it exogenously determined. He argues his approach follows that of Wicksell, who constructed a pure credit economy that is consistent with an endogenous money approach.

“All money is simply someone’s promise to pay. If there is in fact no hard metallic money in which payment can be made, all that can be promised is the convertibility of one promise into another promise ...” (Moore 1985, p. 7)

Bank loans are made primarily at the initiative of the borrower, especially where credit lines have been previously negotiated. Once banks have met credit demands, they hold assets that are traditionally illiquid (today, widespread securitization increases liquidity). Thus, “banks ordinarily cannot increase or reduce their loan portfolios at their initiative.” (Moore 1985, p. 11) Since a reduction in the money supply (bank liabilities) requires a reduction of bank assets, “the ability of central banks to effectuate an exogenous reduction in the money stock will thus depend critically on the composition of bank portfolios.” (op. cit., p. 11) By the 1960s, banks had reduced holdings of reserves and government bonds, making it difficult for them to liquidate positions. Due to private information involved in making loans, banks cannot normally sell them in secondary markets. (Securitization has made this easier—for example, banks no longer hold many long-term home mortgages but instead securitize them.) As a result, “the monetary authorities are rendered relatively impotent in their ability to restrict the rate of growth of the money stock.” (op. cit., p. 12) Thus, “since central banks, as consistent with their paramount supportive role to the financial system, operate to allow the money stock to accommodate increases in the demand for bank credit, monetary aggregates may properly be considered to move endogenously.” (op. cit., p. 24)

Moore argues that banks are price setters and quantity takers both in retail loan markets and retail deposit markets. Bank loans are made at the initiative of borrowers at an interest rate set by banks. Banks also set a deposit interest rate and then passively accept deposits at that rate. Since for individual banks the quantity of loans may exceed the quantity of deposits, they rely on wholesale sources and uses of funds, where banks are price takers. (With the development of large denomination certificates of deposit, banks participate as suppliers and demanders of wholesale funds.) Ultimately, the central bank is the marginal supplier of wholesale funds (at the discount window or in open market purchases) and, as such, the central bank ultimately determines the wholesale cost of funds. Retail loan rates are then set as a markup over the wholesale cost. The retail deposit rate is a markdown of the loan rate, to cover the costs of administering deposits and to provide profits. (Moore 1988b, Chapter 3)

According to Moore, the supply of money should be viewed as horizontal at the going interest rate. That rate, in turn, is ultimately determined by the Fed’s discount rate. The Fed is able to exert only loose control over the supply of money by regulating short-term interest rates by determining the price it charges for reserves. The Fed is not able to constrain bank reserves since banks meet the demand for loans and then must meet legal reserve requirements.

“Reserves are the result, rather than the cause of movements in deposits.” (Moore 1984, p. 107) The interest rate determined by the Fed can affect the demand for loans, but once loans are made, the Fed must supply the reserves needed by the banking system. Moore quotes the Fed’s Holmes affirmatively:

Since banks have to meet their reserve requirements each week ... and since they can do nothing within that week to affect required reserves, that total amount of reserves has to be available to the banking system. The Federal Reserve does have discretion as to how the banks can acquire this predetermined level of needed reserves. The reserves can be supplied from the combination of open market operations and the movement of other reserve factors, or they can come from member bank borrowing at the discount window ... [T]he suggestion that open market operations should be used in the short run to prevent a rise in total reserves through member bank borrowing is completely illogical. Within a statement week, the reserves have to be there; and, in one way or another the Federal Reserve will have to accommodate the need for them. (Holmes 1969, p. 73)

Moore argues that we should view the supply of reserves as essentially horizontal at the central bank’s target rate (fed funds rate in the US). Since the Global Financial Crisis, some central banks have adopted a corridor system in which the market rate is allowed to fluctuate between a ceiling rate (discount rate) and a floor rate (interest paid by the central bank on reserves). The Fed overwhelmingly supplies reserves through open market operations (and, as we will see in Chapter 8, budget deficits generate reserves that are usually removed as the central bank and treasury sell bonds) rather than at the discount window—which is primarily used for necessitous borrowing by troubled banks. The major central banks have moved from a practice of keeping reserves relatively scarce (to prevent interest rates from falling to the floor) to one of maintaining huge excess reserves in the system. In the absence of paying interest on reserves, the overnight rate would fall to zero. Today, the central bank’s control obviously comes through its interest rate target—and not through the quantity of reserves as banks cannot be reserve-constrained when there are massive excess reserves available. This is no longer controversial, at least among heterodox economists.

Moore’s presentation is very similar to Kaldor’s: banks are never quantity constrained, and the money supply function is horizontal at the going rate of interest (which, in turn, is determined by the central bank). We will argue in the next chapter that this is an oversimplification. We might instead presume that, at a point in time, the money supply curve slopes somewhat upward beyond some point because banks require rising interest rates to induce them to increase leverage ratios (the ratio of loans to equity, and the ratio of loans to safe assets such as government bonds), while innovations cause the curve to shift outward over time. Banks are not passive lenders and do not simply

accommodate loan demand. They may not be willing to lend in the face of necessitous borrowing—when there is a run to liquidity and borrowers need to “capitalize” interest (borrow to make payments on their debt—what Minsky calls “Ponzi” finance) during a collapse of expectations. Indeed, banks may also try to shore up their own liquidity—dumping assets at fire sale prices to reduce risk.

When economic prospects look unfavorable, banks revise desired leverage ratios downward and become unwilling to extend credit on demand. Wolfson (1986) argues that banks cut off credit from all but established customers. Banks may experience difficulty in meeting the demand for liquidity by even established customers because a general run to liquidity makes it difficult for banks to maintain a rate of reflux, increases costs of retaining deposits, and reinforces pessimism. Thus, in a recession, the quantity of credit is not demand determined. (Note, this is similar to Marx’s claim that in a crisis, the demand for money reaches its peak—but will not be met by bankers.) None of this, however, conflicts with Moore’s main argument that central bank control is limited to the interest rate target. Indeed, the experiment with Quantitative Easing after the Global Financial Crisis demonstrated that injecting massive excess reserves into the banking system did not cause banks to increase lending. Whatever impact the policy might have had, it came through impacts on interest rates.

Another objection to Moore’s horizontalist approach is that it relies on modern financial arrangements and practices. Indeed, Moore has argued for a “stages” approach to money, according to which the money supply has only become endogenous with the development of modern central bank practices. This implies that the money supply would not be endogenous in the absence of central banks (i.e., the USA before 1913) and says nothing regarding the determination of interest rates in the absence of central banks. However, as we have seen in Chapter 3, the money supply has always been endogenously determined. Indeed, it is more obviously endogenous in the monetary systems that predate the development of a central bank—when banks issued notes as they made loans.

Finally, while Moore’s claim that the central bank sets the overnight interest rate is no longer contested, it leaves determination of other rates up in the air. Moore has argued that Keynes’s liquidity preference theory relied on a presumption of an exogenous money supply and thus must be dropped. As we will see in Chapter 7, however, it is possible to reconcile an endogenous money approach with liquidity preference theory.

Let’s turn to the final contributions to heterodox endogenous money approaches: the Franco-Italian circuit approach, or more generally, the “money as a flow” approach as opposed to the orthodox “money as a stock” approach.

LAVOIE: THE FLOW OF CREDIT AND THE CIRCUIT APPROACH

Lavoie (1984) argues that the major distinction between orthodoxy and the endogenous money approach is that neoclassical theory emphasizes money as a stock, or as an asset in a portfolio, whose value is derived from its usefulness in facilitating the exchange of goods that have already been produced. In Post Keynesian theory, however, it is the flow of money (over time) that is important. Any flow of production requires a flow of new credit or a renewal of past credit flows. Money cannot be introduced into a system after production has taken place, since money creation has to do with the production process itself. In this scheme, borrowers are the causal factor, and the function of banks is to create money that makes production possible. The money stock (measured at a point in time) merely represents credits that have not yet been repaid. (Lavoie 1984, pp. 771–4)

Lavoie classifies those who adopt the endogenous money approach into four groupings: first, Cambridge Post Keynesians such as Robinson, Kaldor, and Kahn; second, American Post Keynesians such as Davidson, Minsky, Weintraub, Moore, and Wojnilower; third, Circuitistes represented by Schmitt, Barrere, and Parguez; and last, Overdraft economists such as Le Bourva, Berger, and Weymuller. The latter two schools are primarily French, and little of their work had been published in English. (Lavoie 1985, pp. 65–6) These four groups accept two ideas. First, loans make deposits, and second, deposits make reserves. Thus, “banks do not wait for the appropriate amount of resources to exist to provide new loans to the public ...” (op. cit., pp. 67–8) Furthermore, as Moore argues, “the Federal Reserve System cannot, any more than the Banque de France, control the emission of high-powered money. It can only modify its discount rate or its rate of intervention.” (op. cit., p. 70) Both ideas have been dealt with sufficiently above (but note that after the Global Financial Crisis, all schools accept that while central banks cannot supply fewer reserves than banks need, they can supply more reserves than banks want so long as the central bank will accept a zero overnight interest rate, or will pay interest on reserves).

There are several points of disagreement among the four schools, however. Lavoie analyzes two points of disagreement between the Overdraft economists and the others. “Overdraft economists believe that there are countries, such as the United States or Germany, where orthodox theory can be applied,” while the endogenous approach applies to countries like France where the central bank accommodates the demand for loans. (op. cit., p. 71) The other groups argue that the approach applies to all economies with sophisticated financial systems and capitalist relations.

Secondly, Overdraft economists believe it is possible to have an economy without credit money and that this may even be preferable. The existence of credit may push interest rates too low to induce saving, so that investment must be financed on the basis of credit. Thus, Overdraft economists believe that the loanable funds theory would be applicable to an economy without credit and that such an economy is possible. This is inconsistent with the endogenous money approach and is also inconsistent with the Keynesian argument that investment determines saving. Credit is seen as an anomaly by the Overdraft economists. But, following Keynes, the expansion of investment requires credit. Terzi (1986–87) and Wray (1988) have shown that saving cannot be a net source of funds for rising investment. Credit is required to fund increases in net investment, which then creates saving. In an economy without privately issued credit money, the government would have to inject credit to allow for net investment. In reality, private institutions provide credit in all developed capitalist economies. Therefore, we should exclude the Overdraft economists from the endogenous money approach, which emphasizes the necessity of credit and money.

The Circuitists and both groups of Post Keynesians agree that it is the flow of credit, not the stock of money, that is important. (Lavoie 1985, p. 74) “They recognize that expenditures must eventually be financed by the transfer of funds into demand deposits. But the stock of these funds at any point in time is considered to be a residual” (op. cit., p. 73) In other words, an expansion of spending requires the creation of credit, which represents the liabilities of some economic units. For example, a firm issues an IOU to a bank and receives a demand deposit that is used to finance an increase in spending. The demand deposit is the liability of the bank and is normally counted as one of the components of the money supply. The quantity of demand deposits outstanding at a point in time is partially determined by the rate at which the firm retires its IOU. Lavoie argues that the important decision takes place when the firm decides to enter into debt in order to increase spending. The firm’s subsequent decision to retire the debt is not considered to be nearly so important. Chapter 9 will demonstrate these points using balance sheets.

Lavoie focuses on the flow of credit rather than on the stock of some monetary aggregate. Thus, his approach is within the endogenous money approach in which money is endogenously created as debt commitments are negotiated. The flow of finance is equivalent to a creation of debt accepted in payment for spending. However, Lavoie may go too far in his assertion that stocks are unimportant. While central banks determine the overnight rate, they do not set the rates across the term structure nor the risk-adjusted rates. Money stocks in portfolios are important because they can affect interest rates. Indeed, Lavoie’s rejection of the importance of stocks probably accounts for his critique of

Minsky's financial instability hypothesis, which he argues borders on a loanable funds approach. We will turn to an analysis of this in Chapter 7.

INTEGRATION OF THE ENDOGENOUS AND STATE MONEY APPROACHES

As we have seen, since the late 1950s there has been a revival of the endogenous money approach that had been followed by the Banking School of the early 19th century as well as by Marx and then Schumpeter several decades later, followed in the postwar period by Gurley and Shaw as well as Minsky. This culminated in the "horizontalist" approach to money advanced by Moore that emphasizes the nondiscretionary nature of reserves. This effectively reverses the "deposit multiplier" of the money and banking textbooks, arguing that "loans make deposits" and "deposits make reserves." The focus is on the private decisions made by banks and their customers, which determine the supply of loans and deposits, hence, the supply of credit money that "endogenously" expands to meet the needs of trade. The central bank can only "exogenously" set the short-term interest rate (fed funds in the US, bank rate in the UK) at which it supplies reserves "horizontally" on demand to banks. Finally, the endogenous money approach rejects even long-run neutrality of money, arguing that the creation of money is tied to the fundamental processes of a capitalist economy; thus, money *always* matters. These points are all well-established in the Post Keynesian literature and do not require further elaboration here. (See Moore 1988b, Goodhart 1989a, and Wray 1990.)

Links of the endogenous money approach to the credit money approach discussed earlier should be obvious. The difference is really one of emphasis, with the endogenous money approach focusing more on bank and central bank decision-making, while Innes was more interested in identifying the nature of credit. The fundamental property of all credit, according to Innes, is that its issuer must accept it in redemption; the endogenous money equivalent is the Banking School reflux principle: excess bank deposits return to banks to retire loans.

What has largely been neglected by some proponents of the endogenous money literature is the role of the state and the impact of fiscal operations on banks and the central bank. At least in the horizontalist version, the role of the state is limited to the central bank's ability to set the overnight interest rate, which requires that it passively accommodate bank demand for reserves. Proponents of the Circuit approach generally limit their analysis to the role of the central bank—and like Moore, usually ignore the impact of fiscal operations on reserves and the money supply. (Exceptions are Bougrine and Seccareccia 2002, and Parguez 2002.) Hence, relations between the state money approach and the endogenous money approach remain relatively less clear.

Some of those who adopt the endogenous money approach have argued that central bank behavior should be analyzed separately from fiscal operations, as the central bank is formally independent of the treasury in many nations. In reality, however—as we will argue in more detail in Chapters 8 and 9—the central bank’s desire to set and hit overnight rate targets means that it cannot be independent of the treasury, in the sense that any undesired impact of fiscal operations on banking system reserves must be immediately and completely offset by central bank operations (except where the central bank is willing to leave excess reserves in the system, letting the overnight rate fall to zero or to the rate paid on reserves).

It is somewhat surprising that fiscal effects on reserves are largely ignored in this revival of the endogenous money approach because these are potentially many times larger than the quantities of reserves directly added or drained by central bank operations. As we’ll see, all else being equal, treasury spending leads to a credit to banking system reserves while tax payments lead to a debit; thus, treasury deficits lead to net credits. In practice, daily operations of the treasury almost always generate either net credits or net debits even if the budget were balanced over the course of the year, for the simple reason that tax payments on any given day would differ from government spending. Hence, the treasury and central bank have created complex procedures to allow them to closely coordinate activities to minimize effects on reserves. These reserve effects of fiscal operations have been a central concern of the MMT literature. (See the discussion in Chapter 8 as well as in Bell 2000, and Bell and Wray 2002–03.)

However, the horizontalist, or more generally the endogenous, money approach is consistent with both the credit and state money approaches. According to the state money approach, the state chooses the unit of account (the dollar) in which the privately issued credit moneys are denominated. The state also chooses which moneys it will accept in payment of taxes and other obligations. In modern sovereign nations with their own domestic, floating currencies, this is always nonconvertible, high-powered money (usually liabilities of the central bank). Private banks help in clearing between the government and the private sector since most taxes are paid through bank accounts and most recipients of treasury checks deposit them into private banks. The central bank and treasury then coordinate activities to offset undesired impacts on bank reserves, allowing the central bank to exogenously set and hit the overnight rate target—as discussed in Chapter 8. The high-powered money accepted by the state is a credit money, a liability of the treasury or central bank, and hence operates according to Innes’s fundamental law of credit—or the law of reflux cited by the followers of the endogenous money approach: state liabilities are redeemed when they return to the state, mostly in tax payments or bond purchases by the non-government sector.

Still, we are left with a question: *must* an endogenous money approach adopt a Chartalist, or state money, approach? An endogenous money approach could possibly locate the origins of money in barter or in a primordial “free market” economy free from government intrusion (as in Parguez and Seccareccia 2000).³ Perhaps it is not so important to finally uncover the “true” origins of money, but rather to explicate the “nature” of money. The question is whether a hypothesized stateless but monetary economy sheds any important light on the nature of the modern money we use. Does it help us to understand the “horizontal” monetary policy of modern central banks? Does it reveal the social relations important to the operation of modern capitalist economies? Is our understanding enhanced by presuming that the money of account originated from the spontaneous adoption by individuals of a numeraire to denominate their private credits and debits?

The view taken here is that the notion that private credit monies are denominated in a Chartalist, national unit of account is more illuminating. Further, there is a *pyramidal* monetary hierarchy that rules clearing—most private clearing of accounts takes place on the balance sheets of banks, while banks use the central bank for ultimate, or net, clearing, all in terms of the national currency. Finally, the central bank is the ultimate clearer for transactions between the private sector and the government. While all credit money—even that issued by the treasury and central bank—represents a liability, this does not put all liabilities on an equal playing field. Further, as most liabilities in any sovereign nation are denominated in the national currency, it becomes possible to clear with liabilities high in the pyramid. If such a hierarchy is important, then starting analysis without a treasury or central bank is deficient.

As Innes argues, the fundamental credit principle is that the issuer—whether household, firm, or government—must accept its own liabilities. Today, only the *sovereign* government(s) can impose liabilities on others.⁴ This puts the state in a privileged position because it can create a demand for its own liabilities by requiring that taxpayers deliver government liabilities in payment of taxes (and other obligations). The nation can also enact legal tender laws and legal reserve requirements to provide further privilege to treasury and central bank liabilities. Finally, the modern state is, of course, a large entity—hence an important purchaser of output and source of income—which makes its liabilities ubiquitous. Still, if the state did not impose obligations payable in its currency and require ultimate payments to itself in the form of its treasury and central bank liabilities, it is difficult to believe that its sheer size and its legal tender laws alone would be sufficient to *guarantee* its current spot at the top of the money hierarchy.

Perhaps the more significant and contentious point concerns the implication of the state money approach for government budgetary issues—the topic of Chapter 8. As Lerner (1943, 1947) recognized, the “money as a creature of the

state” approach leads logically to a “functional finance” view of state budgeting. Because the state spends by emitting its own liability, it does not need tax revenue or the proceeds from borrowing in order to spend. We will return to these points later, but it is important to note how the state money approach conflicts with the conventional “government budget constraint” (GBC) notion, according to which state spending must be “financed” by tax revenues, borrowing, or “printing money.” The constraints faced by a government that spends its own currency are different from those faced by households and firms—a point we take up in Chapter 8.

In any case, there is nothing within the horizontalist or endogenous money approaches that requires a conventional approach to government finance. Indeed, it is a fundamental proposition of horizontalism that central bank provision of its own liabilities is limited only by demand—it can potentially supply reserves without limit. The state money approach simply extends that to the treasury—so long as state liabilities are demanded, they can be supplied by the state (central bank plus treasury—with the division of responsibilities determined by law and procedures developed). As we will see, Lerner’s “creature of the state” and “functional finance” approaches are consistent with the endogenous money approach that insists that reserves are nondiscretionary; indeed, the second principle of functional finance can be seen as a corollary of horizontalism—that bond sales are undertaken to redeem excess reserves. Thus, bond sales, even by the treasury, are nondiscretionary and help the central bank to hit interest rate targets.

Let us conclude for now that the state money and endogenous money approaches can be synthesized. To put it as simply as possible, the state chooses the unit of account in which the various moneys will be denominated. In all modern economies, it does this when it chooses the unit in which payments to the state will be denominated and names what is accepted in payment. The state then issues HPM (high powered money, or base) in its own payments—in the modern economy by crediting bank reserves, and banks, in turn, credit accounts of their depositors. Tax payments lead to debits of deposits from bank accounts and debits to the reserves of those banks. The public and private *giros* are linked through HPM—the state’s money.

Most transactions that do not involve the government take place on the basis of credits and debits, that is, privately-issued credit money. This can be thought of as a leveraging of HPM—since ultimate clearing takes place in HPM. However, this should not be taken the wrong way—there is no fixed leverage ratio (as in the orthodox deposit multiplier story). Further, in all modern monetary systems, the central bank targets an overnight interest rate, standing by to supply HPM on demand (“horizontally”) to the banking sector (or to withdraw it from the banking sector when excess reserves exist—unless it pays interest on excess reserves) to hit its target. In any case, the central bank and

treasury coordinate to ensure the banking system has the reserves required/desired, using bond sales and purchases to adjust reserves to allow the central bank to hit its rate target.

As discussed, there is an important hierarchy in the debt/credit system, with power—especially in the form of command over society’s resources—underlying and deriving from the hierarchy. The ability to impose liabilities, name the unit of account, and issue the money used to pay down these liabilities gives a substantial measure of power to the authority. There is, thus, the potential to use this power to further the social good, although misunderstanding or mystification of the nature of money results in an outcome that is normally below what is economically feasible as the government is “constrained” by the principles of “sound finance.”

Far from springing from the minds of atomistic utility maximizers, money is a social creation. The private credit system leverages state money, which in turn is supported by the state’s ability to impose social obligations, mostly in the form of taxes. Much of the public production is undertaken by emitting state money through government purchases. Much private sector activity, in turn, takes the form of “monetary production,” or M-C-M’ as Marx put it, that is, through the monetary purchase of required inputs with a view to realizing “more money” from the sale of the final product. The initial and final purchases are mostly financed on the basis of credits and debits—that is, “private” money creation.

Because money is fundamental to these production processes, it cannot be neutral. Indeed, it contributes to the creation and evolution of a “logic” to the operation of a capitalist system—discussed in Chapter 10—“disembedding” the economy to a degree never before encountered. (Heilbroner 1985; Wray 1990 p. 54) At the same time, many of the social relations can be, and are, hidden behind a veil of money. This becomes most problematic with respect to misunderstandings about government budgets, where the monetary veil conceals the potential to use the monetary system in the public interest—a topic for Chapter 8.

NOTES

1. As Kalecki (1954) shows, aggregate gross profits are identically equal to the sum of investment, consumption out of profits, the government deficit, and net exports, less saving out of wages.
2. For a discussion of the banking process in Minsky’s theory, see, for example, the section entitled “The Business of Banking” in Minsky’s chapter, “Banking in a Capitalist Economy” in Minsky (1986). Indeed, much of Minsky’s work from 1957 to the present can be seen as an attempt to develop a sophisticated

model of banking. In particular, see Minsky (1957), (1975, Chapter 6), and (1986).

3. See Lau and Smithin for a critique: “the monetary order is socially constructed, rather than deriving automatically from the market. Essentially ... there would have to be a social consensus that the liabilities of the powerful institution ‘count as’ money. It is easy to see, therefore, why historically/empirically the state with its coercive powers has been in the driver’s seat.” (Lau and Smithin 2002, p. 19)
4. In addition to the national government, state and local governments might have such power; in the case of the US, states do have sovereign power to impose obligations—as well as other sovereign powers—and they have also delegated powers to local governments. However, they do not—usually—create their own currency. In the USA the Constitution gives this right only to Congress; however, states might be able to issue negotiable “tax credits” or “state warrants”—as California did in a budget crisis during the GFC. They could be deposited in banks and used to pay taxes and make other payments. This would be a currency in all but name.

7. Money and interest rates

INTRODUCTION

In orthodox theory, the interest rate is determined by the supply and demand of loanable funds. The supply comes from saving, while investment creates the demand. It is believed that banks intermediate between saving and investment, with a higher interest rate attracting more saving. Equilibrium between the downward-sloping investment demand and the upward-sloping saving supply curves determines the interest rate.

Keynes rejected loanable funds theory, arguing that in his theory, the interest rate is determined by liquidity preference. In general, Post-Keynesians still accept the liquidity preference theory of the rate of interest. However, at various points in the previous chapter, we have encountered alternative views on the determination of interest rates. Mainstream Keynesians adopted a money demand-money supply approach to determine equilibrium, although it relies on an exogenous money supply—and Keynes did sometimes argue as if the money supply is fixed—however, this is rejected by proponents of the endogenous money approach adopted in this book.¹ Some endogenous money proponents believe that the liquidity preference theory of interest rates cannot be reconciled with endogenous money. Furthermore, we have also seen that central bank policy has changed significantly since the 1980s (when the horizontalist approach was formulated), with explicit interest rate targeting rather than targeting reserves or the money supply.

In the following sections, we examine the case for a reconciliation of the endogenous money approach with liquidity preference theory and will examine both the *orthodox* Keynesian approach to interest rates and the markup approach taken by some Post Keynesians. We will question some of the propositions of the horizontalist approach—in particular, the horizontal money supply curve and the expectations theory of interest rates. Finally, we will look at how orthodox monetary policy has changed over the last few decades.

THE POST KEYNESIAN CRITIQUE OF THE LOANABLE FUNDS APPROACH

In Chapter 5, we examined Keynes's distinction between finance and saving—indeed, he argued that consumption is a better source of finance than saving is. Both Post Keynesians and Institutionalists also insist that saving does not finance investment; rather, they argue that investment spending creates saving flows. (Foster 1981) In the expanded model (with a government, but ignoring the external sector), investment plus government spending injections determine leakages into saving and taxes; or to put it another way, investment plus the budget deficit equals and *causes* saving. There is thus a link between endogenous money and the recognition that saving does not finance investment; instead, expansion of loans finances spending (of any type).

Keynes rejected the loanable funds theory of interest when he argued that investment determines saving through changes in income:

(T)he initial novelty lies in my maintaining that it is not the rate of interest, but the level of incomes which ensures the equality between saving and investment. (Keynes 1937, p. 250)

(N)o amount of actual investment, however great, can exhaust and exceed the supply of savings, which will always exactly keep pace. (op.cit., p. 248)

Indeed, he argued that saving and investment are “merely different names for the same phenomenon looked at from different points of view.” (Keynes 1987, p. 551) Furthermore, the interest rate cannot be determined by loanable funds, since “the supply curve of savings and the demand curve for investments have no determinant point of intersection, since they lie along one another ...” (ibid., p. 552) As discussed in Chapter 5, they are equal no matter what the interest rate is. Keynes went on to argue that while investment can never be constrained by lack of saving, it can be constrained by lack of finance. However, by finance Keynes did not mean loans. He argued that firms may attempt to accumulate a cash balance before undertaking investment. Since saving can rise only due to investment, this “extra finance involved will constitute an additional demand for money” if investment is increasing. (Keynes 1937, p. 247) Keynes called this the “finance motive” for holding money, and it applies to all categories of planned expenditures. That is, money held to satisfy the finance motive is a function of planned expenditures or aggregate demand. (Davidson 1965) This will be discussed in greater detail below.

What is perhaps less commonly accepted by Post Keynesians is that just as saving cannot finance investment, neither can taxes finance government spending nor can saving finance a budget deficit because the injections must

precede the leakages. These are logical statements of causation. Careful analysis of balance sheets strengthens the argument that taxes cannot provide a prior source of finance for government spending—as shown in Chapter 9. While both investment and saving take place on the balance sheets of the private sector—in the form of private credit money—both government spending and taxes must involve the balance sheets of the government sector (treasury and central bank). It is not possible for taxpayers and their banks to clear accounts with the government in high-powered money (HPM, reserves and cash) unless there is a mechanism for advance provision of the HPM to be used in clearing.

Government spending is the primary source of HPM, although the government (treasury, central bank, or other agents of government) can also lend HPM, purchase assets from the private sector, or provide HPM through transfer payments. Hence, Lerner's functional finance and money-as-a-creature-of-the-state approaches are consistent with the usual Keynesian and Institutionalist views of the temporal relation between injections and leakages: the state must spend (or lend) its liabilities into existence before they can be redeemed in taxation.

What this means is that government deficits do not “absorb” saving—they create it. Further, government bond sales come logically after government spending—creating the saving that can be held in the form of the bonds. What is known as Davidson's “m” (related to liquidity preference—see below) determines what portion of the savings created will be held in bonds; the rest of the bonds must be taken up by financial institutions.

The loanable funds theory claims that an increase in investment or the government's deficit will place upward pressure on interest rates as they compete for a limited supply of saving. As rates rise, the quantity of saving increases. Thus, all else equal, a budget deficit causes interest rates to rise. However, the liquidity preference approach of Keynes and his followers would conclude that deficits increase saving—with no pressure on interest rates, unless liquidity preference were to rise. This, in turn, could affect Davidson's “m”—the propensity to buy bonds. By analogy, the same conclusion would be reached for an open economy context—net exports (an injection) would also cause saving to rise by an equivalent amount.

LIQUIDITY PREFERENCE IN KEYNESIAN AND POST KEYNESIAN THOUGHT

Basil Moore (1988b) argues that Keynes adopted an exogenously determined money supply in the *General Theory*, even though he had previously recognized in the *Treatise* that money is bank debt whose quantity is endogenously determined. According to Moore, Keynes's exposition in the *General Theory* was fatally hampered by his determination to vanquish loanable funds theory

with liquidity preference theory. Furthermore, according to Moore, the logic of the *General Theory* is circular because the demand for money and level of interest rates vary with income, so each change of income would change interest rates and therefore affect investment and lead to further changes in income and interest rates. (Moore 1988c) Thus, the interest rate is indeterminate in the *General Theory*, and was made determinate only by Hicks's IS-LM analysis (where the interest rate is simultaneously determined by equilibrium in goods, money, and bonds markets).

According to Moore, if Keynes had adopted an endogenously determined money supply and exogenously determined interest rates, the arguments would have been strengthened, and Keynes could have defeated loanable funds theory without relying on the "flawed" liquidity preference theory. Lavoie (1985) also rejected liquidity preference theory and argued that Post-Keynesians are confused regarding the determination of the interest rate.² If the money supply is truly endogenous, supplied on demand, then a rise in the demand for finance cannot place pressure on the interest rate. He argued that Keynes's liquidity preference theory applies to an exogenously determined money supply and "Post-Keynesians do not abide by it anymore ..." (Lavoie 1985, p. 76) He also went on to argue that Minsky's financial instability hypothesis "borders on the lack-of-savings thesis and is quite contrary to both post-Keynesian monetary theory and Keynes' view of the trade cycle." (op. cit., p. 77)

We will see, however, that liquidity preference theory is an essential element of both Keynes's theory and the Post Keynesian approach. (Wray 1992) First, the central role played by liquidity preference theory in the general theories of Keynes, Minsky, and the Post-Keynesians is discussed. The following section will show how liquidity preference theory can be incorporated within an endogenous money approach. All of this is important to understanding the *value of money*—the topic of Chapter 10.

Moore's argument that Keynes's liquidity preference theory is "circular" and Lavoie's belief that Post Keynesians have rejected liquidity preference theory are perplexing. Kregel (1980) has argued that the interest rate on money measures liquidity preference, which is the differential between forward and spot liquidity. An own rate of interest can be calculated for each commodity (here, *commodity* is defined very broadly to include anything that can be produced and held through time), which is the return in terms of the commodity on a loan in that commodity. In other words, it is the amount of the commodity that can be bought for forward delivery in terms of a given amount of the same commodity for spot delivery.

The return of a commodity can also be calculated in terms of money, rather than in terms of the commodity itself. This return is called the *marginal efficiency* of the commodity. In equilibrium, the marginal efficiencies of all commodities should be equalized—if the marginal efficiency of one commodity

were higher than that of others, demand would induce production until its expected rate of return fell into line with the marginal efficiencies of all other commodities. Keynes argued:

It seems then that the rate of interest on money plays a peculiar part in setting a limit to the level of employment, since it sets a standard to which the marginal efficiency of a capital-asset must attain if it is to be newly produced ... For it is the greatest of the own-rates of interest ... which rules the roost (because it is the greatest of these rates that the marginal efficiency of a capital-asset must attain if it is to be newly produced); and ... there are reasons why it is the money-rate of interest which is often the greatest ... (Keynes 1964, pp. 222–3)

In the *General Theory*, the *marginal efficiency of capital* determines the spot price of real productive assets, while liquidity preference determines the spot price of financial assets. Monetary factors determine the prices of debts (or the differential between spot and forward money), which then discount the expected stream of profits that would arise from owning real assets (that is, the marginal efficiency of capital). Liquidity preference determines the structure of liability prices, which gives the interest rates and simultaneously determines the demand prices of capital assets. This affects the decision to hold real assets versus financial assets, thereby influencing the level of investment. (Kregel 1987)

Minsky (1986) describes this as Keynes's *two-price system*: one for current output and one for assets. These are connected by the investment decision: capital has a price that comes out of the current output system (since it must be produced) and one that comes out of the asset price system (as capital is a real asset to be used through time). The current output price system generates the *supply price for capital assets* (how much one must pay to induce production) while the *demand price for the capital asset* is determined as described by Kregel. Only if the demand price exceeds the supply price will capital be produced. This forms the basis for Minsky's financial instability hypothesis, which is beyond our present scope.

Kregel has argued that liquidity preference theory and the expenditure multiplier are "two sides of the same coin", thus, "[t]here can be no difference between Keynes's insistence that it is the level of income via the multiplier that equates savings and investment and his insistence that it is the degree of liquidity preference which determines the rate of interest." (Kregel 1988, p. 239). A decline in liquidity preference will lower the interest rate (*own rate on money*), which raises the demand price of capital assets and causes investment to rise until the marginal efficiencies of real assets fall to equality with the lower interest rate. This is equivalent to arguing that income rises through the investment multiplier until saving rises to equality with the new higher level of

income. This is a *state of rest*—the Keynesian definition of equilibrium, where there is no incentive to change behavior.

Keynes's theory can be seen as a liquidity preference theory of nominal asset prices or a *theory of value* for assets: a decrease in liquidity preference increases the relative money prices of capital (and other illiquid) assets, generating profit opportunities, inducing investment, and raising income through the multiplier. (Kregel 1988) Again, we will use these arguments in Chapter 10 to develop a theory of value of money.

KEYNES ON MONEY SUPPLY AND MONEY DEMAND

Keynes carefully distinguished his theory of the determination of interest rates from that of Ohlin, Robertson, and Hicks, who argued that the interest rate is determined by the intersection of the demand and supply of credit or loans. He argued that their view “is exactly the same as the classical doctrine, over again, to the effect that the quantity of saving depends on the rate of interest.” (Keynes 1937, p. 244) Their view is concerned “with changes in the demand for bank borrowing, whereas I am concerned with changes in the demand for money.” (Keynes 1937, p. 245) What Keynes meant is that the demand and supply of loans in a given monetary circuit cannot place pressure on the interest rate as long as bank loans advanced to finance spending immediately return to banks. Only if bank advances leak into cash hoards would there be pressure on interest rates. Given a degree of liquidity preference of the public and given the willingness of banks to satisfy the demand for hoards, the interest rate will be established without reference to the demand and supply of loans. However, if there is an increase in planned expenditure, this may be accompanied by an increase in hoards to satisfy the finance motive unless banks accommodate the rising demand for hoards. (See Chapter 5.)

Thus, as Kregel (1986) argued, more investment is subject only to a liquidity constraint: the banking system must allow its balance sheet to adjust by satisfying the additional demand for hoards at the prevailing interest rate. Normally, banks will advance credit to meet an increase in the demand for money used to finance investment expenditures with no pressure on interest rates. However, there will be upward pressure on interest rates if, first, the liquidity preference of the public rises so that bank credit does not return to banks, and second, if banks meet the extra money demand only at higher interest rates. Keynes argued:

[T]oo great a press of uncompleted investment decisions is quite capable of exhausting the available finance, if the banking system is unwilling to increase the supply of money and the supply from existing holders is inelastic. (Keynes 1973, p. 210)

As Chick (1983, p. 240) argues, consumption and investment spending, as well as profit and wage income, primarily remain within the banking system as bank liabilities are used to circulate such flows as a “revolving fund” of finance. The revolving fund itself would finance investment with no pressure on interest rates. (Keynes 1973, p. 209) An increase in investment (or any other kind of spending) would require a larger revolving fund and would have to be met by banks. But only an increase in demand for hoards or decrease in willingness to supply hoards would affect interest rates. The supply of liquidity would depend on the terms “on which the banking system is ready to become more or less *unliquid*.” (Keynes 1973, p. 219; emphasis added)

While Keynes focused on the relation between the finance motive and planned investment, Davidson (1978) extended the analysis to include other components of aggregate demand. He defines the transactions demand for money as a hoard that is held in anticipation of spending to satisfy the finance motive—and not as money that is held as some proportion of income. Keynes suggested that it is possible an increase in planned government expenditures might be associated with increased hoards if the government borrowed money in anticipation of future spending. (In Chapter 8, we cast doubt on this proposition.) Similarly, planned increases in consumption or exports could also increase this transactions demand for money. Furthermore, planned shifts of spending from one component to another might increase demand for money even in the absence of an increase in total spending. For example, if consumers have lower access to credit than do firms, a planned shift from investment to consumption spending would increase the need for transactions balances. However, both Keynes and Davidson suggest that banks normally would satisfy this through overdraft facilities.

To briefly summarize Keynes’s argument: The demand for hoards (determined by psychological propensities and the finance motive) and the supply of hoards (determined by the willingness of banks and others to provide money now for more money later) together determine the interest rate. “If the banks stand firm, an increased propensity to hoard raises the rate of interest, and thereby lowers the prices of capital assets other than cash, until people give up the idea of selling them or of refraining from buying them in order to increase their hoards.” (Keynes 1937, p. 251)

Some Post Keynesians have extended Keynes’s argument to broaden the range of assets that can satisfy the desire for liquidity. Thus, “hoards” might include short-term liquid assets in addition to cash. (Dow and Dow 1989) Mott argues “liquidity preference is a theory of the desire to hold short- versus long-term assets.” (Mott 1985–86, p. 230) While rising investment must generate sufficient saving to cover it, if liquidity preference is high, households prefer to hold some of their saving in liquid form, rather than purchase long-term bonds. As mentioned above, Davidson has defined “*m*” as the marginal propensity to

buy long-term bonds out of personal saving. (Davidson 1978, p. 299) If m is less than one, then even though investment generates saving, some of it will be held in liquid form so that some of the long-term bonds issued to finance investment cannot be sold to households. In this case, a higher interest rate on long-term assets relative to that on short-term assets is required to induce households to take up bonds issued to fund investment projects (Davidson 1986). Otherwise, financial institutions will be forced to hold the bonds so that households can remain liquid.

As Kahn argued, the “quantity of money is the means by which the public hold that part of their wealth which is looked after by the banking system.” (Kahn 1954, p. 238) In other words, the banks can take up the long-term instruments and allow the public to hold short-term debts (“money”). The prices of long-term assets (or the long-term interest rate) must be at a level that ensures the public will hold those securities that the banking system will not. (Kahn 1954, p. 238) Thus, if we take m as an inverse function of the liquidity premium (the differential between interest rates on short-term and long-term assets), high liquidity preference means that it would take a prohibitively high liquidity premium to raise m close to unity, so that the banking system is required to remain illiquid in order that the public can hold liquid assets.

Liquidity preference has also been related to expected profitability: as profit expectations increase, the liquidity premium required to induce households to relinquish hoards falls. For example, Mott argues, “the state of liquidity preference is governed by the profitability of business.” (Mott 1985–86, p. 231) Robinson argues that if profit expectations rise, this will encourage more investment and if the banking system meets the increased needs of trade by expanding its balance sheet, there is no rise in interest rates. In fact, one expects that this would be the case, as bankers should also be subject to increased optimism. Thus, “the improved prospect of profit counts twice over - once in promoting investment at a given cost of borrowing and once in lowering the cost of borrowing.” (Robinson 1979, p. 157)

On the other hand, a collapse of profit expectations will probably be associated with a rising liquidity preference. The public as well as firms will want to shift out of long-term assets and into liquid assets, while banks will be reluctant to supply short-term liabilities and hold long-term bonds. Thus, low profit expectations count “twice over”: they reduce the incentive to spend on investment goods, and the willingness of banks to provide credit at the existing interest rate. Thus, we expect that rising liquidity preference and falling profit expectations will be associated with rising interest rates.

All of this is consistent with the view of Gurley and Shaw cited above (in Chapter 6): increased spending is financed through the expansion of credit. There is no reason to expect that an increase in the supply of short-term credit should immediately cause interest rates to rise, since this supply is largely

demand determined. However, as firms increase production and income is generated, surplus sectors accumulate net wealth. Gurley and Shaw argue that the surplus units may want to diversify portfolios. Thus, a liquidity premium is required to induce surplus units to hold net wealth in the form of long-term bonds. Alternatively, surplus units can hold the liquid liabilities of financial intermediaries, and financial institutions can take up the long-term bonds—as discussed. In any case, continued growth of (debt-financed) investment may eventually push up interest rates due to liquidity preference and the diversification motive.

Finally, liquidity preference determines the premium required to induce surplus units to hold long-term assets. This then sets the standard to which the marginal efficiencies of all assets must conform. Thus, liquidity preference goes into the determination of the prices of assets and thereby affects investment and aggregate demand, in general. There is no reason to expect the system to generate full employment because there is no guarantee that the interest rate set by liquidity preference will be low enough to generate sufficient investment. The alternative way of looking at it is that the marginal efficiency of investment is too low relative to the marginal efficiency of money determined by spot and forward prices. This also explains why money can't be neutral—it goes into determining asset prices, the level of investment, and the state of effective demand.

In conclusion, liquidity preference theory is consistent with the endogenous money approach. If we recognize modern institutional practices—such as business credit lines, consumer credit cards, and the way government really spends (see Chapter 8)—the exposition of the role played by a “finance motive” and the propensity to hoard cash by both Keynes and some of his early Post Keynesian followers is probably overwrought. Households, firms, and governments do not normally increase hoards substantially in advance of planned spending.

We should, instead, focus on the role played by liquidity preference in determining asset prices and the level of investment. With modern central banks acting as lenders of last resort, and with deposit insurance standing behind banks, we do not observe runs on *insured* bank deposits, nor runs into cash.³ However, there are runs to liquidity—such as the one that triggered the Global Financial Crisis, when a wide swath of financial assets that had been considered highly liquid suddenly became illiquid—so their prices collapsed. More recently, in the USA, there was a run out of uninsured deposits at banks such as Silicon Valley Bank that was triggered by the Fed's rate hikes that sharply lowered the market price of what should have been very safe assets—US Treasury bonds and highly rated mortgage-backed securities—issued in a low-interest-rate environment. Suddenly, only the Fed could restore liquidity to those markets by lending against troubled assets at face value. While the protected part

of the financial system is relatively immune to runs, the unprotected part is larger and still vulnerable.

FLOWS, STOCKS, MONEY DEMAND, AND LIQUIDITY PREFERENCE

In Keynes's theory, interest rates are closely related to stocks—assets that can be held—although if demand prices of producible assets are high enough relative to supply prices, more can be produced. A great deal of confusion has resulted because his critics (and some of his followers) have failed to distinguish consistently between stocks and flows and between liquidity preference and the demand for credit. Investment and saving are flows, and investment expenditures necessarily generate equivalent saving flows. The finance of investment flows does not directly affect interest rates, just as saving flows cannot directly influence interest rates. Keynes argues:

Credit, in the sense of “finance”, looks after a flow of investment. It is a revolving fund which can be used over and over again ... The same “finance” can tackle one investment after another. (Keynes 1973, p. 209)

A revolving fund of a given size can continuously finance a given flow of investment and saving⁴ However, saving and investment flows have effects on stocks. Spending on investment that is financed by debt will, at the same time, provide long-term assets that can be accumulated by savers. As discussed earlier, since we normally expect that the marginal propensity to purchase long-term bonds out of increments to personal saving is less than one, portfolio preferences will require a premium to induce the public to hold the long-term assets issued by firms to finance positions in the capital assets. Note that the premium is not due to a lack of saving—the investment spending has created equivalent saving. The public could save until it is “blue in the face” (as Keynes put it), and this would not eliminate the premium. It is the preference for liquid assets that necessitates a premium to clear the long-term bond market.

Often the terms liquidity preference and money demand are used interchangeably. However, a distinction can be made between them. Money demand can be defined as a willingness to issue debt to finance spending, linking it primarily to expenditure flows—as is common in the Circuitist approach and also in the Horizontalist endogenous money approach—which argues that an increase in money demand is met by an increase in money supplied through loans. Money demand thus defined does not directly impact interest rates. However, an increase in money demand that is met by the expansion of bank balance sheets will affect balance sheet (stock) positions. In this way, money demand might indirectly affect interest rates, as will be discussed below. On

the other hand, liquidity preference is defined as a desire to exchange illiquid items on a balance sheet for more liquid items, or even to decrease the size of a balance sheet by retiring debt. Liquidity preference is thus related to stocks and directly influences interest rates.

In an economy with developed financial institutions, deposit insurance, and a central bank that is willing to act as a lender of last resort, hoards of cash or narrow money are relatively unimportant. Furthermore, lines of credit or overdraft facilities can satisfy the “finance motive” without requiring cash hoards. Thus, Keynes’s notion that the interest rate is determined by the supply and demand of hoards does not strictly apply to developed capitalist economies. However, if liquidity preference is restated as “a theory of the desire to hold short- versus long-term assets,” then it is consistent with the analysis presented above. (Mott 1985–86, p. 230) Banks typically provide firms with short-term finance that is retired when an investment project is complete and long-term bonds are floated. Investment spending generates an equivalent amount of saving, received primarily as bank deposits by surplus units. It is common to use the term *saving* as the flow created by investment, and *savings* as the accumulated stock of deposits.

We have argued that liquidity preference can be measured as the interest rate differential required to induce surplus units to exchange bank deposits for long-term bonds (or as the interest rate differential required to induce banks to purchase long-term bonds while issuing shorter-term liabilities). Following Keynes, we should reject a flow determination of the interest rate and instead locate its determination in the decision to hold stocks. This does not imply that flows do not have an indirect impact on interest rates, since flows affect stocks. An expansion of bank balance sheets might affect interest rates because banks are concerned with their balance sheet position (a stock). Similarly, if firms and households decide to shift their portfolio to liquid positions, this will also affect interest rates, which must adjust so that households will become satisfied with the existing amount of liquidity. This may not be a smooth adjustment process—it could generate a panic if wealth holders realize that liquidity cannot be obtained. The important point is that liquidity preference determines interest rates, and not money demand.

As Dow and Dow (1989) argue, the supply of credit and liquidity preference are interdependent. When liquidity preference is low, banks meet the demand for credit by expanding their balance sheets. As liquidity preference rises, banks become increasingly unwilling to meet the demand for credit, so that credit rationing replaces accommodative behavior. “The overall picture, then, is of credit creation accommodating demand when liquidity preference is low, but not when liquidity preference is high.” (Dow and Dow 1989, p. 161)

In summary, liquidity preference is measured by the premium required to induce wealth holders to accept illiquid portfolios. Because an increase in the

size of a bank's balance sheet can be associated with increasing leverage (of capital and safe liquid assets), beyond some point, banks will raise interest rates as an expansion proceeds. This tendency for interest rates to rise as externally financed spending increases can be offset by many factors, including:

- (1) a shift of depositor preferences toward illiquid positions (that is, liquidity preference falls);
- (2) competition by other (non-bank) financial institutions;
- (3) central bank provision of reserves through open market purchases of less liquid assets; and
- (4) optimistic expectations concerning the future (so that liquidity preference falls as profit expectations rise).

Thus, liquidity preference theory is not inconsistent with the endogenous money approach. Investment is not constrained by the supply of saving, nor is it constrained by the quantity of money (narrowly defined). As banks are not strictly quantity constrained, their willingness to make loans must be a function of their willingness to take less liquid positions in assets. This willingness will surely be affected by the state of long-term expectations regarding profitability. However, it will also be affected by the balance sheet positions of banks, the portfolio preferences of the public, and the policies of the central bank.

Liquidity preference theory is not simply a "flawed" expedient adopted in a critique of loanable funds theory. It is a necessary part of Keynes's general theory. Liquidity preference theory is closely tied to Keynes's theories of effective demand, of the relation between the marginal efficiency of capital and investment, and of the finance motive. Indeed, liquidity preference theory strengthens the endogenous money approach, just as the endogenous money approach strengthens Keynes's general theory.

ENDOGENOUS MONEY, EXPECTATIONS THEORY, AND THE MARKUP APPROACH

Some of those who have adopted the endogenous money approach have replaced liquidity preference theory with a markup theory of interest rates that incorporates expectations in long-term interest rates. In this section, we critically examine that view.

Moore (1988a, 1988b) relates the fed funds rate to the discount rate set by the Fed. The Fed can set it within a range (the size of the range depends on the actions of other central banks), and bank arbitrage then keeps all other short-term interest rates in close alignment. The interest rates charged by banks are a markup over costs, which include costs of wholesale funds, labor costs, and

profit on equity invested in the banks. The long-term interest rates are then set by arbitrage to equalize the expected holding period yields, so that the long-term interest rate is determined by current expectations of the level of future short-term interest rates that will be set through central bank policy.

Like Moore (1988b), Rouseas (1985) argues that banks are price setters but quantity takers in retail loan and deposit markets.⁵ On the other hand, they are quantity setters and price takers in markets for wholesale sources of funds (such as large denomination certificates of deposit). The interest rate charged on bank loans is a markup over the cost of wholesale liabilities issued by banks. Rouseas refers to Kalecki's oligopoly theory of the markup in which the degree of market power governs the markup that can be added to prime costs. He takes the fed funds rate as an indication of the prime cost of obtaining funds and argues that the prime rate charged on business loans is a more-or-less stable markup over the fed funds rate.

Several objections have been raised to arguments such as these. Robinson (1979, pp. 145–8) criticized the view that long-term interest rates are determined by the expected value of future short-term rates, arguing that it makes just as much sense to argue that the long-term rates determine the short-term rates. Keynes argued that the long-term interest rate is determined by speculation on the long-term interest rate itself, rather than by speculation on the future course of short-term interest rates (as Hicks and Kaldor had argued). (Kahn 1954, p. 229) Kahn explains that expectations about the short-term rate will influence the long-term rate only to the extent that they affect expectations about the long-term rate. (Kahn 1954, p. 235) If tight money policy forces up short-term rates, and if people expect higher short-term rates to continue far into the future, then expectations about the long-term rate will probably also shift upward. This, according to Kahn, “does not really amount to saying more than that the long-term rate of interest is influenced by expectations of banking policy.” (Kahn 1954, p. 234)

He goes on to give an example in which the expectations theory provides an illogical conclusion. Assume banks buy short-term bills from the public and sell long-term bonds. The short-term interest rate will tend to fall as prices rise, while the long-term interest rate will tend to rise as prices fall. According to the expectations approach, the bond rate rises because the expected future short-term rates are rising even though the actual short-term rate is falling. (Kahn 1954, p. 235) Alternatively, assume that the banks buy long-term bonds, causing the long-term rate of interest to fall. This fall in the long-term rate has nothing to do with expectations about the future short-term rate. Indeed, the current short-term rate might be rising if the banks are financing their bond purchases by selling short-term bills. On the other hand, if banks financed bond purchases by issuing deposits, there would be no pressure. Simple expectations theory is not a very useful guide in such cases.

Expectations theory also has trouble if decisions can be reversed. As Kahn argues, “a decision to go long rather than short, or vice versa, is not indissoluble.” (Kahn 1954, p. 236) The purchase of a short-term bill does not indicate that the buyer will remain in bills. When one goes short, the only relevant bond rate is the one that will prevail when the short-term bill matures. While that expectation will be related to the expected bill rate on the same date, and while the bond rate on that date will also be related to the bond rate on future dates as well as to expected future bill rates, this does not mean that the long-term rate is merely the expected value of future short-term rates.

The course of the price of a bond is not necessarily relevant even for a bondholder. (Note that it could be relevant for banks if they are forced by regulators to continually mark to market.) The relevant price is that which will reign on the day on which the bond will be sold. One is not normally forced to sell bonds when a capital loss would be incurred. (Chick 1983, p. 215) If one buys a bond with the intention of selling at the end of a year, this could indicate a belief that on that particular date, the long-term rate on the bond will be above the short-term rate that will prevail. This is the position of those who adopt the expectations theory. However, it is also possible that one buys the bond because of the belief that its rate is greater than the rate likely to prevail one year hence on long-term bonds. This was Keynes’s position.⁶

For these reasons, the expectations theory of long-term rates seems to be insufficient.

ENDOGENOUS MONEY: INFINITELY ELASTIC OR CREDIT-RATIONED?

The markup theory takes Fed behavior as the determinant of the discount rate that determines the Fed funds rate (or some other wholesale interest rate), which is the major component of prime costs. We can view the supply of reserves as horizontal at that rate. These are reasonable assumptions. Given that wholesale cost of funds, banks add a markup to establish the retail loan rate of interest. According to Horizontalists, the money supply curve is then horizontal (infinitely elastic) at that rate of interest, with banks meeting all demand for loans at that rate.

We will argue that this is another oversimplification if banks face uncertainty over the performance of the economy and impacts on borrowers’ ability to pay. In that case, we should not view the money supply curve as horizontal at a given rate of interest. Banks are concerned with the ratio of loans to safe assets, such as government bonds plus reserves, and with the ratio of loans to equity. Should borrowers default on loans, the liquidation of safe assets permits banks to meet payment commitments temporarily, while equity allows banks to absorb loan losses. The money supply curve is not horizontal even if

all created deposits return to the banking system, because banks must be concerned with rising leverage ratios. Even where additional loans can be made that are no more risky than loans made previously, banks may require higher interest rates to compensate for riskier leverage ratios. Let's examine in more detail the argument against a horizontal loan supply.

Because uncertainty must always exist, it is desirable to establish close relations with customers. Minsky has said that the first rule of banking is that customers lie. That is, customers will try to place their financial position in a favorable light. According to Minsky, the second rule of banking is that bankers lie, too. That is, banks also try to present their position favorably to calm the fears of depositors and stock holders. Even if neither bankers nor their borrowers lied, banking is a risky business because the future is uncertain and money commitments are necessarily future-oriented. For these reasons, banks use a combination of price and quantity constraints to ration credit.

A substantial literature has developed to provide explanations for quantity rationing by banks, largely by New Keynesians. Calomiris, et. al., (1986) examine the credit market for small farmers, who typically borrow from small, local banks. They find, "information-intensive localized customer borrowing relationships, rather than impersonal debt and equity markets, are important in agricultural finance." (Calomiris, et. al., 1986, p. 443) In such markets, prices do not ration credit. Bankers fear that if they raise interest rates too high, farmers who are good credit risks will go elsewhere, while those who are bad risks will remain. Thus, banks attempt to establish close relationships with local farmers to screen potential defaulters. They use quantity rather than price to allocate credit. Blinder and Stiglitz (1983) apply a similar analysis to non-agricultural loans. They argue that capital markets are not perfect: some borrowers may default, and it is difficult or impossible to gain perfect information regarding risks. Banks specialize in acquiring information about default risk, but this information is firm-specific and difficult to sell. Banks devise non-price screening mechanisms because rising interest rates may decrease expected returns due to adverse effects on the mix of applicants (less credit-worthy applicants are more willing to pay the higher interest rates, particularly if they expect to default).

Similarly, Stiglitz and Weiss (1981) argue that banks use quantity rationing rather than price rationing because higher interest rates would lead to adverse selection, in which only the riskier borrowers would accept loans. Furthermore, higher interest rates would encourage borrowers to take on riskier projects with the hope of earning higher gross profits to meet debt commitments (adverse incentive). Thus, quantity rationing can be consistent with maximizing behavior.

Bernanke (1981, 1983) argues that borrowers and lenders develop relationships to lower the costs of credit intermediation. He argues that bank failures

(such as those that occurred during 1930–33) raise the cost of credit intermediation by eliminating these special relationships. Thus, potential borrowers (whose bank has failed) are credit constrained until they are able to develop a new relationship with another bank. While they are credit constrained, their spending declines, which exacerbates the disruption due to aggregate demand effects.

Finally, Dow and Saville (1988) argue that banks are risk averse and face uncertainty concerning loan defaults. They do not know which interest rate would adequately compensate for the risk of default, since such risks are not known with precision. Because banks pool risks, they are partially protected from random loan defaults. However, they can further reduce potential losses by placing quantitative limits on each customer. Although some part of the loan can be made against collateral, the typical loan will be made against expected future income. Thus, the bank limits the loan amount to some portion of secured property and expected income. The bank will also provide credit only to customers for whom credit worthiness has been established. This assessment cannot be based only on objective considerations but must also be subjectively determined by intimate knowledge of the customer.

Like Minsky, Dow and Saville argue that bank perceptions of appropriate leverage ratios can change over the business cycle. Higher income levels encourage banks to make loans based on greater debt-to-equity ratios. Revisions of rules of thumb permit greater leverage of prospective income flows. Bank managers follow only general rules, but each tries to keep in step with other managers. Overly rapid growth of a particular bank would raise suspicion about the quality of the bank's loans. This could raise the cost of issuing liabilities and might lead to a run on uninsured deposits.

Banks also follow similar pricing schedules—partly due to competitive pressures, but also to avoid raising eyebrows. A bank that pays more for deposits will be thought to be experiencing trouble attracting deposits, while one that charges less on loans will be thought excessively risky. As each bank must maintain the confidence of other banks (which are potential sources of inter-bank loans), none can deviate substantially from norms of behavior without attracting unwanted attention. However, innovations and revisions of norms do occur and can greatly affect the price and quantity of debt acceptable to banks.

These considerations strengthen the argument that credit constraints are normally operative, but the norms can change over a cycle or on a longer trend. An individual bank that has been growing more rapidly than others might refuse credit or demand a high interest rate to demonstrate that it is prudent. The prospective borrower may then be credit constrained and may not be able to go to another bank since banks require information about creditworthiness—and this is costly and requires time to obtain, as Bernanke argued. Thus, micro considerations lead to macro credit constraints. This would be particularly true

as expectations fall in a recession, during which individual banks have an even greater incentive to avoid growing faster than others. This is because banks that grow faster than average will lose reserves through the clearing mechanism and will face escalating costs of recapturing them as expectations decline and liquidity preference rises, forcing interest rates to rise.

Dow and Saville conclude that interest rates and the quantity of loans are primarily determined by norms of behavior. (Dow and Saville 1988, p. 61) Speculators influence interest rates, each trying to outguess average opinion of the rate that norms will establish. However, because expectations are loosely held and because the constraints on possible interest rates are rather loose, markets like to follow a leader. For this reason, the actions of the central bank have a larger influence than one would guess by looking at the relative size of central bank interventions. If the central bank raises the discount rate, this is likely to affect expectations and cause other interest rates to rise. If markets come to expect tight money policy toward the peak of an expansion, interest rates might rise even before the central bank announces higher interest rates. Furthermore, since expectations are loosely held, changes in a few key interest rates (such as the certificate of deposit rate or loan rate set at one of the money center banks) will likely have an influence over expectations. Therefore, even though outstanding stocks of assets and liabilities are large relative to additions to bank balance sheets during any given period, there is no reason to suppose that actions of banks have little effect on interest rates. For this reason, the liquidity preference of banks will play a role in the determination of expectations held by speculators.

Recent changes in central bank behavior and financial innovations have combined to reduce the importance of bank rationing, however. In the US, the Fed and other regulators have deregulated financial markets, reducing geographical, functional, and institutional barriers to entry. This has increased competition for the funds of surplus units and increased competition to fund the spending of deficit units. Financial markets have come to operate more like auction markets and less like markets based on rationing and established relations. (De Cecco 1987, p. 6) Increased competition reduced loan rates of interest, expanded secondary markets for long-term bonds (and reduced their interest rates), and increased credit availability for those who previously had been credit constrained.

Moreover, this literature seems dated after the development of securitization, which supposedly reduces all these loan risks through diversification and tranching, and moves debt off the balance sheets of the originating banks. Further, much of the literature was based on the highly suspect assumption that borrowers have more information about their individual risks than banks do (the bank that makes commercial loans to local restaurants likely has much better information on default risk than does the optimistic owner who

is opening a new restaurant). However, the significance of such considerations is that the interest rate is not set to clear the market. There is always a “fringe of unsatisfied borrowers”—those who receive no credit or less than desired. This means that the money supply curve is not horizontal and that banks are not “price setters and quantity takers” in retail loan markets. While Moore notes that the typical credit line is only half-utilized, this does not mean that banks have met all credit demands. Credit lines are not equally distributed. As Rousseas (1986, p. 65) argues, most unused overdrafts remain unused as they are not where they are most needed—and the interest rates on overdrafts can be penalizing.

Wojnilower argues that the trend toward securitization and auction markets would have been quickly reversed in the absence of central bank guarantees. (Wojnilower 1987) That is, as soon as one or two spectacular failures led to default on securitized loans, surplus units would have quickly returned to the safe, short-term liabilities of commercial banks, and banks would have returned to established relations with customers. However, because the central bank steps in to prevent default, the new institutional practices are validated (in Minsky’s terminology) so that the new instruments retain their liquidity. The public came to view uninsured certificates of deposit or uninsured money market funds as nearly perfect substitutes for insured demand deposits. As a result, banks must compete on price, paying high deposit rates of interest and purchasing risky assets in an attempt to earn profits. De Cecco argues that this trend necessarily increased instability.

As we now know, the pace of innovation increased, and assessment of risk declined during the late 1990s and early 2000s as financial markets came to expect that successive Fed Chairs Greenspan and Bernanke would come to the rescue. And when the financial system really did crash into the Great Recession after 2007, the Fed came in with \$29 trillion in lending and spending on securities to save the global financial system. (Felkerson 2011)

CENTRAL BANK POLICY, LIQUIDITY PREFERENCE, AND INTEREST RATES ON LONGER MATURITIES

When Basil Moore developed the horizontalist approach in the early 1980s, central banks proclaimed that they were targeting the money supply by controlling reserves. Moore argued that they could not do so—central banks always accommodate the demand for reserves, so their policy instrument is the interest rate at which they supply them. By the early 1990s, the major central banks had abandoned reserve and money targets and admitted that they actually target interest rates, as Moore had claimed they were doing.

Today, the Fed and other central banks announce their interest rate target (fed funds rate in the USA). A lower target will place downward pressure on a

variety of interest rates. This may, in turn, lower loan interest rates of interest if a given differential (loan rate minus deposit rate) is maintained. (The degree of competition is one of the factors that determines whether the differential can increase.) If the central bank raises the target rate, this will tend to increase competition for other sources of funds, which will tend to raise deposit and loan rates of interest.

However, the central bank should consider the impact of rate hikes on bank costs of issuing liabilities—which tends to rise more quickly than the rate they earn on assets. The interest rate differential must be sufficient to allow banks to earn the normal rate of profit on bank equity. Restrictive central bank policy will gradually push up both the deposit and loan rates of interest. However, a sharp increase in the overnight rate by the central bank can cause severe problems by raising the costs of issuing short-term liabilities faster than banks can increase their earnings on longer-term assets.⁷ This is why the Fed switched in the 1990s to using a long series of very small rate changes (called “gradualism”) as it tightens or loosens policy. Still, with that caveat, we conclude that short-term interest rates are under the control of the central bank.

What about other interest rates? Moore adopted the markup approach: loan rates are set as a markup over the cost of reserves (fed funds rate) and deposit rates are a markdown below the central bank’s target. The markup approach implies that Keynes was mistaken in his belief that profit-seeking behavior in a monetary economy subject to uncertainty necessarily leads to unemployment and instability. If liquidity preference plays no role in the determination of interest rates, then the central bank could achieve full employment by announcing that the interest rate would be held at some low and constant rate forever—indeed, this was a position that Moore sometimes defended.

However, it appears to be inconsistent with Keynesianism. The markup approach does not provide a theory explaining how dynamic forces generate unemployment and financial instability. The essential insights of Keynes (and of Minsky—the financial instability hypothesis) are left out of the markup approach. Keynes argued that a monetary economy inherently generates an interest rate that is chronically too high to establish a level of effective demand consistent with full employment. In the markup approach, the authorities need only lower interest rates to eliminate unemployment, which is not caused by any intrinsic features of capitalism but by bad policy.

Keynes held out little hope that an easy money policy would generate low enough interest rates to maintain full employment. The general case, according to Keynes, is a situation in which expected effective demand and expected profits would lead to the equalization of the marginal efficiency of investment with the interest rate at a level of investment too low to achieve full employment. Moore reduces this to a special case where the central bank has set the

interest rate too high (he sometimes expressed bewilderment that Japan's decades of near-zero rates did not generate economic growth).

If liquidity preference is inversely related to profit expectations, then low expectations cannot be overcome by monetary policy—no matter how easy it is. High liquidity preference prevents easy monetary policy from pushing the marginal efficiency of money below the marginal efficiency of capital. Thus, banking behavior, as well as the behavior of the public, prevents enlightened policy from succeeding. As we saw in the aftermath of the Global Financial Crisis (GFC), zero interest rate policy (ZIRP) plus trillions of dollars, pounds, euros, and yen spent during multiple phases of Quantitative Easing (QE—central bank purchases of bonds) did not boost loans for spending.

According to Keynes, if the expected returns to capital investment are low (or even negative), then even a very low interest rate will not generate investment. He argued that “beyond a certain point money's yield from liquidity does not fall in response to an increase in its quantity to anything approaching the extent to which the yield from other types of assets falls when their quantity is comparably increased.” (Keynes 1964, p. 233). The reasons for this include the high liquidity premium and low carrying cost of money, as well as zero elasticity of production, employment, and substitution of money.⁸ While central banks did lower their target rates to zero, this didn't spur borrowing because expectations of recovery remained pessimistic—so expected returns from investing were below the cost of finance (which remained well above target rates).

While it is true that credit money can be created essentially on demand, Keynes emphasized that high liquidity preference cannot be curtailed by favorable lending terms. Moore equates liquidity preference with money demand—that is, a demand for loans. But a rise in liquidity preference is not the same as an increase in money demand. As Marx put it, the demand for money might reach its peak in a financial crisis, but this is not a demand for credit to finance spending; rather, it is a demand for money to repay debts.

Lavoie and Seccareccia (1988) point out that the central bank may not be able to push down interest rates on longer maturity debt when liquidity preference has risen because speculators are likely to believe that interest rates will eventually rise again. Moore (1988c) argues that easy money policy during the Great Depression pushed the treasury bill rate in the US down to 0.02 percent, but the long-term bond rate remained above 2.5 percent. He argues that this is a measure of the markup over the discount rate. The discount rate was 5.19 percent in 1929, fell to 2.6 percent in 1933, and reached a minimum of 1 percent by 1939. In contrast, the Baa rate on bonds was 5.60 percent in 1929, rose to 7.76 percent in 1933, and fell only to 4.96 percent by 1939. These data seem inconsistent with a markup approach, since the markup increased from 0.41 percentage points in 1929 to 5.16 in 1933.⁹ However, it is consistent with

a liquidity preference approach since a collapse of profit expectations during the depression should have been accompanied by rising liquidity preference. It also confirms Keynes's belief that "easy money" policy would not be sufficient to push down interest rates sufficiently to achieve full employment when expectations are pessimistic.

Keynes argued that due to interest rate (or "capital") risk on longer maturity bonds, it is probably impossible to push longer rates below about 2 percent because the potential capital losses outweigh the interest earned. Modern financial theory uses a bond's duration to assess that risk. While there are several methods of calculating the risk, they lead to a similar conclusion: long duration (which is related to term to maturity) means more interest rate risk. Future central bank policy is therefore an important consideration when deciding whether to buy assets with longer maturities, although so is credit risk. And, unfortunately, the future is uncertain, including future central bank policy. However, in Keynes's view, it is fluctuations of the marginal efficiency of investment rather than fluctuations of interest rates (at least in part driven by central bank policy) that dominate economic performance. Liquidity preference goes into determining both lending rates and the marginal efficiency of investment. So rising liquidity raises the cost of funds and lowers the expected return from investing.

In conclusion, while the central bank can push the target rate to zero, and while this might reduce the interest rate on some assets (such as government bills) close to zero, this does not mean the loan rate of interest or the private bond rate of interest will decline commensurately—or that banks will lend to all who are willing to pay the loan rate. Banks will price and ration credit to customers based on established relationships, rules of thumb, and liquidity preference. Liquidity preference will still play a role in determining the premium required to induce surplus units to hold bonds that are not backed by the Fed or full faith and credit of the treasury.

THE RISE OF THE NEW MONETARY CONSENSUS AND EXPECTATIONS THEORY

Beginning in the 1990s, a new approach to central banking began to take hold—based on the New Monetary (or Macro) Consensus (NMC)—and it was dominant by the time Chairman Bernanke had taken over the Fed from Alan Greenspan (see Papadimitriou and Wray 2021 for a critique). The good news is that the old exogenous money approach and money targeting were dropped in favor of an interest rate target and an implicit acceptance of endogenous money (generally without acknowledging it). The bad news is that it retains most of the neoclassical assumptions, including, most importantly that the

economy moves toward a market-clearing equilibrium with full employment and only short-run rigidities temporarily preventing equilibrium.

The main monetary policy goal is to hit an inflation target. The NMC propagates the notion that central banks implement policy by controlling inflation expectations that cause inflation and disequilibrium, so central banks need to create the expectation that inflation will remain low. This is done by sending signals using the overnight rate target and standing ready to raise it even before inflation arrives. That signal then causes markets to expect the central bank will continue to raise rates to fight inflation, causing long-term interest rates to rise. That slows the overheated economy and tames inflation.

We will not go deeply into the details of the theory here, but it is based on the view that the economy can move above or below a full market-clearing equilibrium in the short run—called a “demand gap.” A positive demand gap means the economy is overheated; a negative demand gap means it is operating below capacity. Other than random shocks, the main disturbance preventing general equilibrium is a deviation of the real interest rate (the nominal rate adjusted for inflation) from the natural interest rate (the rate consistent with full equilibrium). The policy solution is to adjust the nominal target rate to move the real rate to the natural rate. Both adaptive and rational expectations come into play as inflation expectations that determine the real rate are both backward-looking and forward-looking. Instead of using the money supply as a policy tool, the central bank follows a Taylor Rule in setting the overnight rate target. To the extent that it can align inflation expectations with its own ultimate inflation target, the Fed is able to get actual inflation in its target range (typically close to a measured 2 percent rate). Greenspan argued in 1994 that central bank success at achieving this is the key to ensuring economic growth at a sustainable pace. (See Papadimitriou and Wray 2021)

The NMC approach failed as assessed by the Fed’s ability to control expectations—both in the aftermath of the GFC when even zero rates could not *raise* expectations of inflation nor stimulate lending and spending by *lowering* long-term rates, and in the aftermath of the COVID crisis, when *rising* target rates could not *lower* expectations of inflation (that remained relatively constant) nor *raise* longer-term rates sufficiently to slow spending, even when the Fed kept warning markets that target rates would go much higher and stay high much longer than markets expected. (Papadimitriou and Wray 2021)

Note the similarity of the NMC view and the horizontalist view on the determination of long-term interest rates—that are supposed to be determined by expectations of future central bank settings of the overnight rate. Yet, the Fed’s real-world test of its ability to control long-term rates through expectations failed to do so in the aftermath of the GFC. It is, of course, possible that the markets simply did not believe the Fed, but that seems implausible after the GFC when central banks adopted ZIRP and QE for a decade. The

near-continual ZIRP policy in the dozen years after the GFC did not cause households to raise their expectations of inflation at all, nor did professionals raise theirs. They held their expectations essentially constant over the whole period (with households expecting inflation to be a point higher and professionals expecting inflation to be slightly higher than actual), while the Fed undershot its inflation target nine years in a row.

Former Fed Governor Daniel Tarullo criticized the Fed for its attachment to “problematic concepts and hard-to-estimate variables.” (Tarullo 2017, p. 2) Following the NMC strategy, the central bank is supposed to set its interest rate policy with a forward-looking economic forecast and a view as to how the Fed’s policy might improve the outcome. He worries that the outcome depends on unobservables, and some of those depend on the relation of two series on which the *data* are unobservable—in a sense, doubly unobservable—increasing uncertainty about outcomes. The unobservables important for monetary policy based on the NMC include potential GDP growth, the natural rate of unemployment, and the natural interest rate. While we can look backward at some variables that are supposed to affect these unobservables, for policy-making that does no good, because we must look forward. The fourth critical unobservable is the inflation rate to which the economy might converge; but this is finessed because central bankers presume that it is 2 percent—the typical inflation target they pursue.

Tarullo was particularly critical of the overreliance on inflation expectations, noting they are “arguably a unique” kind of unobservable. While we do have surveys of the inflation expectations of professional economists, consumers, and financial market traders, ironically, we do not have surveys of nonfinancial businesses that have the power to set prices. The others do not. Businesses say they set their prices based on what the market will bear, not according to inflation expectations. Tarullo argues that the Fed does not have a good handle on the mechanism through which expectations are supposed to affect actual inflation.

The Fed presumes—without a well-grounded theory—that it can affect expectations, which will then determine inflation. When inflation is low, it takes credit for “well-anchored inflation expectations” that result in low inflation—the Fed’s focus on fighting inflation was successful because inflation was low. That is tautology. However, as noted, after the GFC, the inflation rate was below target for nearly a decade. After the COVID recovery, inflation and inflation expectations remained stubborn even with rapid increases to the target rate, and with the Fed continually warning it would keep rates high as long as it takes. It is a plausible conclusion that neither the Fed’s actions nor expectations have much to do with inflation outcomes.¹⁰

CONCLUSIONS

The central bank today sets the policy interest rate for overnight funds in reserve markets to achieve a set of policy objectives. Such objectives might include other interest rate targets, bank leverage ratios, rates of growth of bank balance sheets, degree of competition from shadow banks, unemployment rates, exchange rates, inflation, and rate of growth. The transmission of central bank actions to the economy is exceedingly complex and subject to great uncertainty because profit-seeking bankers, other financial institutions, and nonfinancial institutions, as well as consumers, actively subvert policy when it seeks to constrain profitable behavior. As Minsky said, policy works only by changing behavior—and monetary policy may not give much leverage to change it.¹¹ Further, even though the target rate is set by the central bank, that doesn't mean it can ignore economic conditions and possible impacts—which means that policy is constrained.

Central banks have dropped the notion that they determine the money supply and instead openly target the overnight rate. Some central banks now operate with a corridor system—paying interest on reserves to set a floor to the overnight rate, setting a discount rate as a ceiling, and thereby determining the corridor within which the overnight rate remains. Sometimes central banks also intervene into asset markets to constrain rate movements on longer-term assets such as government bonds and mortgage-backed bonds—as with QE. In principle, a central bank can determine the rate on any financial asset by standing ready to buy/sell at announced prices—something they did not seem to understand as they turned to QE, when they tried to influence rates using quantity rather than price (they would announce how many bonds they would buy, rather than setting a price—that would have directly set the interest rate).

Central bank policy does influence bank behavior and does go into the determination of costs faced by banks. However, policy can be swamped by liquidity preference, innovations, competition, and profit expectations. The liquidity preference of banks can affect the loan rate of interest in conjunction with the liquidity preference of the nonbank public, which together determine the bond rate of interest. As bonds (and other forms of non-bank finance) and loans are partial substitutes, competition from shadow banks helps to determine the loan rate of interest. The loan rate of interest goes into the determination of the deposit rate of interest, since it is the differential between these two that largely determines the profitability of banking. Again, nonbank sources of liquid assets compete with banks and help determine the deposit rate of interest, as does the liquidity preference of the public.

Market power, as well as regulations established by law or authority, affect both loan rates and deposit rates. Horizontalists are right to emphasize the

difference between retail and wholesale markets, although rampant securitization, as well as use of credit scores (which are backward-looking) rather than careful underwriting (which is looking forward) has reduced the importance of retail banking. Banks, as well as other kinds of financial institutions, increasingly originate loans to sell them quickly into markets that are often highly speculative, manipulated, and even fraudulent, as we found in the *post-mortem* examination of the causes of the Global Financial Crisis. And market power has become much more concentrated in the hands of a half dozen financial behemoths—at least in the USA, which used to have less concentration (other countries, like Canada, already had high levels of concentration). Banking now looks very different than it did in the 1980s when the endogenous money approach was revived by Post Keynesians.

The theory of interest rates that comes out of an endogenous money approach is neither a simple markup theory nor a theory of the demand and supply of money (as in the ISLM model). The interest rate on loans is largely determined by the willingness of banks and other financial institutions to allow their balance sheets to expand. This willingness is determined by rules of thumb, custom, and expectations. The central bank's interest rate policy is also important, but this policy is constrained by the ability of banks to adjust to higher interest rates in the case of a tight money policy, and by bank willingness to lower interest rates in the case of a loose money policy.

Speculators attempt to obtain arbitrage profits among assets of differing maturities. This helps to maintain a “term structure” of interest rates. However, the term structure is subject to variation because there is no objective basis determining what is the proper structure. Thus, each speculator forms expectations trying to outguess average opinion. As Keynes argued, the interest rate is influenced by what people expect it to be. Since there is little objective basis determining interest rates, they are established by rules of thumb and by norms of behavior—and even by attempts to rig the markets—although with the central bank setting the base rate with its target. Thus, an increase in money demanded to finance rising aggregate demand will not raise interest rates unless financial market participants believe it should. If banks believe larger balance sheets would exceed proper leverage ratios, then rates must rise to compensate for greater perceived risk. If the public believes that the addition of bonds to portfolios would increase risk, then bonds can be floated to finance investment only at rising interest rates.

In other words, both the money supply and interest rates are endogenously determined.

NOTES

1. To some extent, Keynes did oversimplify his approach to money in the *General Theory*—as he had already examined the institutional details in his earlier *Treatise*. As he admitted in the *General Theory*, “whilst it is found that money enters into the economic scheme in an essential and peculiar manner, technical monetary detail falls into the background.” (Keynes 1964, p. vii)
2. During a run to liquidity, changes in the interest rate will not equilibrate liquidity preference with the quantity of liquid assets. Indeed, the types of assets that satisfy liquidity preference during an expansion may not be deemed liquid during a downturn. Thus, not only is liquidity preference determined by expectations, but the liquidity of assets is also subjectively determined, at least in part.
3. These arguments apply to the commercial banking system as a whole. An individual bank that increases short-term loans to provide for working capital expenses can expect to lose deposits through the clearing mechanism. However, this bank will also gain deposits lost by competitors. If a bank grows at the same pace as its competitors, it will not lose deposits through clearing. A bank growing substantially faster than others will be forced to recapture lost deposits through liabilities issued in wholesale markets (such as the Fed funds market). This reflux or redemption need not be instantaneous, for banks generally have several weeks to meet reserve requirements. A rapidly growing bank is likely to face higher costs and will have to accept lower gross profit margins (loan rate of interest minus wholesale cost of funds).
4. In fairness to Rousseas (1986), he does not accept Moore’s “horizontalist” approach. Rousseas argues that the central bank does not fully accommodate, and that the interest rate is not completely exogenously set by the central bank. Tight money policy will raise interest rates, which induce innovations to shift the velocity-interest rate function. This is similar to Minsky (1957). However, Rousseas does seem to argue that liquidity preference theory should be dropped in favor of a markup approach to interest rates.
5. Neither expectations theory nor the markup approach is well supported empirically. Mankiw and Summers (1984) have shown that the evidence is not consistent with the expectations theory of interest rates. Furthermore, the evidence presented by Rousseas (1985, 1986) does not support the hypothesis that the markup of the prime rate over the Fed funds rate is stable. Laudadio (1987) shows that the differential between the interest rate charged on large loans and that charged on small loans moves countercyclically in ways that would suggest that a greater markup is added to large loans during expansions but that a smaller markup is added during recessions, while the markup on small loans moves in the reverse direction. The explanation for this is not immediately obvious.
6. Periodically, central banks do push up interest rates quickly as the Fed did in its 1979 experiment in Monetarism, during which the discount rate rose from 7.46 percent (1978) to 13.42 percent (1981). The experiment disrupted

financial markets and forced the Fed to abandon tight monetary policy. The experiment shows that banks cannot quickly adjust to a higher regime of interest rates by simply adding a markup to the discount rate. However, it does show that rapidly rising interest rates lead to disintermediation and innovation. Eventually, financial institutions can adjust to higher interest rates. The discount rate averaged just over 1 percent during the 1940s, while it averaged more than 7 percent between 1983 and 1987, even though inflation averaged nearly 5.7 percent per year during the 1940s and only 3.34 percent during 1983–87. (All data were taken from *The Economic Report of the President*, 1988. Inflation is the rate of increase of the consumer price index.)

7. Keynes also argued that stickiness of wages in terms of money also enhances money's liquidity premium. In the concluding chapter of the *General Theory*, Keynes argues that the authorities should push the interest rate as low as possible, but believes it is "unlikely that the influence of banking policy on the rate of interest will be sufficient by itself to determine an optimum rate of investment." (Keynes 1964, p. 378)
8. All data are from *The Economic Report of the President*, 1989.
9. See Nersisyan and Wray, <https://www.levyinstitute.org/publications/whats-causing-accelerating-inflation-pandemic-or-policy-response>.
10. See Wray 2007, https://www.levyinstitute.org/pubs/wp_515.pdf.
11. See Wray 2012, <https://www.levyinstitute.org/publications/global-financial-crisis>.

8. Sovereign currency and fiscal policy

INTRODUCTION: CONVENTIONAL VIEWS

According to the conventional view, tax revenue provides the income needed by the government to finance spending. A government might be able to spend in excess of its revenue, at least temporarily, if it is able to issue debt that the public will hold. A sovereign government can also issue non-interest-bearing debt—currency—to finance deficits, but that is almost universally scorned. This can be done directly by the Treasury or indirectly through central bank purchases of government debt. Because money-financed deficits would expand the money supply, many economists claim it would cause inflation following the Monetarist belief.

If, instead, the government sells bonds, the money supply would increase—according to most economists—only if the central bank accommodates by purchasing bonds from banks, thereby increasing bank reserves and the money supply (through the money multiplier). As long as the central bank does not accommodate, there should be no direct impact on inflation since the bond sales would remove reserves and money from the economy. (If the government's spending creates excess aggregate demand, that would create inflation—regardless of the method used to finance it.)¹

However, government borrowing is likely, according to orthodoxy, to *crowd out* private borrowing by adding to the demand for loanable funds, driving up interest rates, and displacing interest-sensitive private spending (investment, housing, consumer durables). (In the ISLM model, the result is similar, but interest rates are driven up by rising transactions demand for money facing a fixed money supply.) Crowding out can be partial, complete, or even more than complete, depending on assumptions. Over the longer run, if crowding out does occur, this will depress aggregate supply and could thereby induce cost-push inflation. Hence, even bond-financed government deficits should be avoided as much as possible. While most economists recognize that at least under some situations government deficits are desirable, most argue that persistent deficits must be avoided.

Even Keynesian economists generally argue that structural deficits should be avoided: while a government should run deficits in recessions, these should be offset by surpluses during expansions. Permanent deficits must be avoided

because no government can operate in a manner that generates the expectation that it will never be able to retire its debt. While there may not be a specific debt-to-GDP or deficit-to-GDP ratio at which markets lose trust in the government, it is not doubted that such thresholds exist. Governments are believed to be subject to market forces that determine the quantity of debt they can safely issue as well as the price (interest rate implied) of the debt. Beyond that limit, the risk of default raises the interest rate the government must pay on its debt.

If the domestic population will not take up all the debt, the government will be forced to sell bonds in international markets. This requires that foreigners are willing to lend to the government running a deficit. Too much borrowing would cause the exchange rate to fall, and the interest rate charged to rise—neither of which is good for the domestic economy. International markets might even force the government to borrow in a foreign currency (issue foreign-currency-denominated debt) if the country's finances are questionable. The government may be forced to impose domestic austerity to placate international markets so it can sell bonds internationally at a reasonable interest rate and without tanking the currency. But austerity lowers the growth rate, and if the interest rate on government debt exceeds the economy's growth rate, then the debt-to-GDP ratio rises without limit.

We do not intend to explore these positions in more detail. Rather, our analysis will demonstrate that this view misunderstands the nature of government spending, taxing, budget deficits, and bond sales. Permanent government deficits are the theoretical and practical norm in a modern economy. While it is possible to run a surplus over a short period, as we will discuss, this has income and balance sheet effects that unleash strong deflationary forces. Given usual private-sector preferences regarding net saving, economic growth requires persistent government deficits—except in the case of a country that can run significant and permanent current account surpluses.

Further, government spending is always financed through the creation of money rather than directly through tax revenues or bond sales. This does not mean that taxes and bond sales are unimportant, but that they primarily serve other purposes. As we have argued earlier, the government's acceptance of money in payment is an important determinant of what will serve as money in an economy. Taxes redeem the government's currency just as loan repayments redeem bank money. Further, tax receipts play a role in current operating procedures adopted by treasuries and central banks that allow spending to proceed. Finally, government bonds have been important for central bank operations and are held as a safe asset in private portfolios.

Understanding "how government really spends" leads to an entirely different view of the degree to which governments are *forced* to respond to pressures coming from financial and international markets. We will argue that most of the pressures that governments currently believe arise from markets

are self-imposed constraints that come from a misunderstanding of the nature of government deficits. In this chapter, we will focus on fiscal operations, but these necessarily include central bank operations too. In Chapter 9, we will dig more deeply into banking and central banking—including interest rate setting, which is normally thought of as monetary policy.

ABBA LERNER AND FUNCTIONAL FINANCE

Our view builds upon the Keynesian approach and is most closely related to Abba Lerner's functional finance approach. According to Lerner,

The central idea is that government fiscal policy, its spending and taxing, its borrowing and repayment of loans, its issue of new money, and its withdrawal of money, shall all be undertaken with an eye only to the results of these actions on the economy and not to any established traditional doctrine about what is sound or unsound. (Lerner 1943, p. 39)

He went on to list two laws of functional finance:

The first financial responsibility of the government (since nobody else can undertake that responsibility) is to keep the total rate of spending in the country on goods and services neither greater nor less than that rate which at the current prices would buy all the goods that it is possible to produce. (*ibid.*)

When spending is too high, the government should reduce spending and raise taxes; when spending is too low, the government should increase spending and lower taxes. He invokes the analogy of steering a car down the road—this is essentially the Keynesian “pump-priming” approach to managing economic growth, using spending and taxes to keep effective demand at the right level.

An interesting corollary is that taxing is never to be undertaken merely because the government needs to make money payments ... Taxation should therefore be imposed only when it is desirable that the taxpayers shall have less money to spend. (*op. cit.*, p. 40)

If the government is not to use taxes to “make money payments,” then how are these to be made? According to Lerner, the government should not always turn to borrowing for the purposes of spending, because “The second law of Functional Finance is that the government should borrow money only if it is desirable that the public should have less money and more government bonds.” (*ibid.*) In other words, the functional purpose of taxes and bonds is not really to finance spending as each serves a different purpose—taxes remove excessive private income while bonds offer an interest-earning alternative to money. Instead, the government should meet its needs “by printing new money”

whenever the first and second principles of functional finance dictate that neither taxes nor bond sales are required. If effective demand will not be driven too high by additional spending, the government should not raise taxes in step with increased spending. And if the interest rate will not be driven too low by “printing new money” to finance the spending, then there is no reason to sell bonds. This is similar to a Keynesian “money supply and money demand” approach to interest rates: drain money by selling bonds to reduce its supply only if the government wants the interest rate to be higher.

In summary, Lerner argued:

Functional Finance rejects completely the traditional doctrines of “sound finance” and the principle of trying to balance the budget over a solar year or any other arbitrary period. In their place it prescribes: first, the adjustment of total spending (by everybody in the economy, including the government) in order to eliminate both unemployment and inflation, using government spending when total spending is too low and taxation when total spending is too high; second, the adjustment of public holdings of money and of government bonds, by government borrowing or debt repayment, in order to achieve the rate of interest which results in the most desirable level of investment; and third, the printing, hoarding or destruction of money as needed for carrying out the first two parts of the program. (op. cit., p. 41)

He concluded that functional finance “is applicable to any society in which money is used as an important element in the economic mechanism.” (op. cit., p. 50)

Some of Lerner’s terminology is problematic, and his “steering wheel” analogy is misleading. Modern government doesn’t really spend by issuing currency, and central banks do not need to manage the money supply to target the interest rate. Today, central banks simply announce interest rate targets to signal to banks that they stand ready to buy or sell bonds, which will be sufficient to move market rates to the target. Furthermore, Lerner’s analogy in which the government is supposed to use spending and taxing to “drive” the economy down the middle of the lane is far too simplistic. Still, the thrust of the goals of functional finance is on the right track.

When one first encounters Lerner’s arguments, they seem quite puzzling. How can the government spend without tax revenue or borrowing? What is the limit—if any—to the government’s spending? What we are adding in this book to Lerner’s approach is (1) an explicit recognition of the role played by taxes; (2) an explicit examination of the impact on reserves of the application of the second principle of functional finance (treated in more detail in Chapter 9 below); and (3) an analysis of a government spending program that will automatically generate full employment—Lerner’s recommended goal (discussed in Chapter 11).

In the next section, we return to Knapp's State Money or Chartalist approach to sovereign currency. Note that Abba Lerner also adopted this view, taking a "money is a creature of the state" approach in another important article. (Lerner 1947) Indeed, the functional finance approach can only apply to a nation that has its own currency. In our final section, we will use the Chartalist approach to examine the case of the euro—a monetary system that was intentionally set up to delink nations from their currencies, which could eliminate the possibility of applying the functional finance approach.

SOVEREIGN CURRENCY

As discussed in Chapter 4, all modern capitalist economies operate with a Chartalist money system (although the original set-up of the Euro area was a deviation from this). This is often called "fiat money" but we will call it a sovereign currency. For reasons explained in Chapter 4, the government cannot simply *declare* that its currency must be accepted in payment, that is, by "fiat." Rather, it needs to "drive" the demand for the currency. As Knapp argued, state "acceptation" in payment determines what will serve as the national (*valuta*) money. The main driver is the obligation the state imposes to make citizens liable for payment of taxes, fees, and fines (and perhaps tribute or tithes in earlier societies) using its own currency.

Let us define what we mean by currency sovereignty. First, the state chooses its own money of account—a point we have made earlier. Second, it issues a currency denominated in that money of account. Currency can take the form of coins, paper notes, and deposits at the central bank. In modern economies, central bank reserves are by far the most important form taken by currency. As we will see, almost all payments to and by government involve central bank reserves. Of course, that has not always been true, as central banking is only a few hundred years old (and the US did not create a permanent central bank until 1913). This leads to the third point: today, payments by and to the government take the form of central bank reserves. However, recipients of government spending generally receive the payment through a credit to their bank accounts. As discussed in Chapter 3, the government and private money *giros* are integrated in the modern financial system.

Fourth, currency sovereignty does not include issuing government obligations denominated in a foreign currency. The governments of rich countries typically do not issue foreign currency-denominated debt, but many others do. Those that do commit their governments to making payments in foreign currency, and—importantly—they may find themselves unable to meet that requirement. So, the principles of functional finance do not apply to government spending or borrowing in a foreign currency. The fifth requirement follows from this: currency sovereignty implies that the country does not peg

the exchange rate. If a country adopts a fixed exchange rate against a foreign currency (typically, the US dollar), then it is effectively guaranteeing that it will deliver the foreign currency. Its debts are not really in its own currency, as it is promising to convert that to a foreign currency at a fixed exchange rate. It might not be able to make good on the promise.

If these conditions *are* met, then the government has a sovereign currency. We will use the term “sovereign government” to indicate that it has *currency* sovereignty. There are other meanings attached to the term *sovereign*, but we are emphasizing currency sovereignty.

Why is currency sovereignty important? We will argue that it ensures greater fiscal and monetary policy space. The most important implication is that affordability—as normally defined—is not an issue. If we take the USA as an example, dollar sovereignty ensures that the national government can afford to buy anything that is for sale in dollars. The federal government cannot “run out of” dollars. This does not mean that the government should spend without limit. It does not mean that the government faces no limits. It does not mean that “deficits don’t matter.” It does not mean that “debt cannot be too big.” But it does mean that unlike any other entity (household, firm, local and state governments), it can make all promised payments as they come due. Involuntary default on commitments in its own currency is not possible. If the government does default on its own sovereign currency commitments, it has done so by choice. We will return to this issue.

The sovereign currency is issued directly by the treasury (as in the US, where the Treasury issues coins) or through the central bank (as in the US, where the Fed issues paper notes as well as reserves that are essentially demand deposits for banks) as a *non-convertible* government liability (non-convertible means the government itself does not promise to convert it at a fixed exchange rate; of course, foreign exchange markets will exchange the currency at the market value). This sovereign currency generally functions as legal tender, that is, it is sanctioned by the courts as the money that fulfills “all debts, public and private.” It is the only money that is ultimately accepted in payment of taxes (a point we return to below). It is also the money into which bank liabilities are generally convertible (either on demand or after some specified waiting period), and that is used for net clearing among banks and between private banks and the central bank.

It is also the money used as the link between the public and private pay communities, as discussed earlier in Chapter 3. It is the money that sits at the top of the debt pyramid (or hierarchy), or the *definitive* and *valuta* money (using Knapp’s terminology). It is the most liquid liability used domestically—except in rare circumstances where a foreign currency is used domestically (as is the case in some of the formerly socialist countries and in some Latin American nations today)—often alongside domestic currency. The most important thing

to understand is that in a normally functioning modern economy, the domestic sovereign currency is always accepted in exchange for domestic production—anything that is for sale in the USA with a dollar price can be had by delivering US currency (coins, notes, or central bank reserves).

Looking at it from the public's side, obligations to the government can be met with the sovereign currency. In the USA, taxpayers pay their tax obligations using US currency. They rarely use coins or notes, so overwhelmingly, taxes are paid with bank reserves held at the Fed. When taxpayers pay the Internal Revenue Service by check or electronically, their deposits are debited, and their banks' reserves are debited. Effectively, their banks make the tax payments for them, using reserves deposited at the Fed. The Fed simply debits the reserve account, and the bank debits the taxpayer's account. As discussed, in modern economies, the public and private *giros* are connected through the central bank. Payments to and by government involve both private banks and the central bank. Private banks make payments for and to households and firms; the central bank makes payments for and to the treasury.

There is an additional important point to make. In addition to voluntary default, a sovereign government can impose barriers to spending on itself. These can take many forms. The most common and reasonable constraint is to adopt a budget and to make it difficult to exceed the budgeted spending. This helps to hold the government accountable and to prevent excessive spending that could cause inflation. Indeed, rather than adopting a requirement to balance the budget, it makes more sense to impose on the budgeting process a requirement that the inflationary aspects of the spending (and taxing) approved in the planned budget must be assessed before budget approval. Inflationary pressures arise not just from the size of the budgeted spending but also depend on the components of the spending. Unfortunately, policymakers usually focus on the proposed budget's impact on the deficit rather than its impact on inflation—an important point to consider later. In our view, that is the biggest lacuna in today's budgeting process.

In addition to a budgeting process, the USA also imposes a debt limit. As we will see, operating procedures typically adopted include the routine issuance of government bonds, and in the USA, bonds are counted in the calculation of debt subject to the debt limit (by contrast, cash and Fed reserves are not counted). Over the course of every year, bonds are continually issued (as discussed below), and if that debt limit established by Congress is reached, the Treasury must ensure it is not exceeded—typically by delaying some payments and prioritizing spending. Historically, Congress has always eventually raised the limit to allow normal spending to resume, although the process has become increasingly politically fraught. In a worst-case scenario, Congress would refuse to raise the limit and the Treasury could be forced to miss payments on debt and other obligations. This would be a technical default. Clearly,

it would be a “voluntary” default, as Congress could avoid it simply by raising the debt limit. While one could argue that political considerations “forced” the default, it would not have been pressed onto the government by external forces, such as refusal by markets to accept more government debt. Default is a political decision, not an economically-imposed event.

Finally, there are two restraints that are very common: central banks are prohibited from purchasing bonds directly from the treasury, and they are prohibited from allowing treasury overdrafts on deposit accounts held at the central bank. Both prohibitions are motivated by a desire to prevent the treasury from what is called “money-financed” deficit spending. The great fear is that if the treasury is allowed to do this, it will cause inflation; further, it will be incentivized to spend “without limit.” We will argue that these operationally imposed constraints, as well as the thinking behind them, are misguided. Budgets constrain spending—the treasury needs budget authority to spend. That is a sufficient constraint on spending. It makes no sense to approve spending during the budgeting process and then prevent the treasury from spending because the outstanding stock of debt has reached some arbitrary limit. Besides, all treasury spending is functionally “money financed”—as is all spending by all types of economic units—meaning that there is no reason to drive an operational wedge between the nation’s central bank and treasury. These points will require more explanation.

HOW DOES A SOVEREIGN GOVERNMENT SPEND?

When a modern government spends, it issues a cheque drawn on the treasury (or, increasingly, through an electronic payment). Its liabilities increase by the amount of the expenditure and its assets increase (in the case of a purchase) or some other liabilities are reduced (in the case of a social transfer—for example, the government’s social security payment liabilities are reduced by the amount of social security checks issued). The recipient of the treasury check will almost certainly “cash” the check at a bank; either the recipient will withdraw currency, or, more commonly, the recipient’s bank account will be credited. In the latter case, bank reserves are credited by the Fed in the amount of the increase of the deposit account. Increasingly, all this is done “electronically”, but that changes nothing of substance.

The treasury payment increases the recipient’s deposit at the bank, and the central bank credits that bank’s reserves. The treasury’s liability (the check) is converted to the central bank’s liability (reserves). At the same time, when the treasury makes a payment, its deposit account at the central bank is debited. To reiterate, the central bank makes and receives payments on behalf of the treasury—just as a household’s bank makes and receives payments on behalf

of the household. There is nothing strange about this. (We will go through the balance sheets in the next chapter.)

For our purposes, it is usually not important to distinguish between the Fed's and the Treasury's balance sheet when discussing sovereign currency. When the Treasury makes a payment, that shows up as an increase in reserves held by a private bank. Those bank reserves carried on books as bank assets and as the Fed's liability are nothing less than a claim on government sovereign currency—at any time, a bank can convert these to coins or paper notes or use them in payments to the state (usually on behalf of depositors, but also to pay their own debts to the central bank).

In the USA, taxes are received throughout the year (although not uniformly, as receipts are concentrated around quarterly dates, as well as the USA's April 15th tax deadline). What this means is that government payments *made* and taxes *received* over any short period are not likely to be matched—even if there is a perfect match over the fiscal year (which would occur if the budget were balanced that year). Until recently, tax payments in the USA were initially received mainly into special tax accounts held by the Treasury at private commercial banks. The Treasury can also place revenues from bond sales into the tax and loan accounts—that is why “loan” is included in the name of the accounts (a bond sale is treated as if it were a loan to the Treasury). Alternatively, receipts from taxes and bond sales can go directly into the Treasury's account at the Fed. Finally, the Treasury can place a “call” on its deposits in the tax and loan accounts to move them to its deposit account at the Fed. (Again, we will use balance sheets in Chapter 9 to demonstrate these points.)

Recall that a typical constraint prohibits a treasury from running an overdraft on its central bank deposit account. To avoid that, the US Treasury would transfer funds from private banks to its account at the Fed when it wished to spend. If the tax and loan accounts did not have sufficient funds, the Treasury would sell bonds to the banks to replenish deposits. Normally, private banks use their reserve balances at the Fed to purchase Treasury bonds; however, banks that offer tax and loan accounts to the Treasury are allowed to purchase them directly by crediting the Treasury's account (again, this is as if they are making a loan to the Treasury). They would lose reserves only when the Treasury places a “call” on the account—to move its deposits to its account at the Fed so that it can spend. Note that when the Treasury spends, the reserves are restored to the banking system as a whole—although they are not necessarily received by the banks that lost reserves. However, the fed funds market allows banks to recapture reserves.

With the advent of Quantitative Easing during the Global Financial Crisis, so many reserves were pumped into the banking system that the Treasury stopped using these accounts (see Chapter 9 below). There was no reason to

postpone debits to bank reserves when taxes were paid because banks had far more reserves than needed. So, the receipts from taxes and bond sales have gone directly into the Treasury's account at the Fed, and payments by the Treasury lead directly to a debit to that account. To ensure sufficient deposits are held in the account at the Fed, the Treasury has been maintaining larger balances. When these begin to run down, the Treasury issues more bonds. As banks have been operating with massive excess reserves since the GFC, they have had plenty of reserves to buy more bonds.

These operations will be used continuously whether the government budget is in balance, in deficit, or in surplus over the course of a year. This is because the daily mismatch between treasury spending and receipt of tax revenues is huge. Any mismatch will affect bank reserves, and timeliness is of the essence—both for banks and for the treasury. The treasury cannot run overdrafts, at least overnight (“daylight” overdrafts certainly occur), so it uses operations such as bond sales to replenish deposits. Banks must have reserves to clear accounts with each other and with the central bank, so the central bank accommodates their needs.

It is important to understand—therefore it bears repeating—that these operations are undertaken regardless of the budgetary outcome over the course of the year. Even if the Treasury runs a balanced budget over a fiscal year, it will sell bonds to deal with the daily mismatch between tax receipts and spending to avoid overdrawing its account at the Fed.

The particular operating procedures adopted for these purposes can vary somewhat across countries and over time. The procedures evolve largely to ensure smooth payments clearing with minimal impact of government spending and taxing on bank reserves, and to minimize treasury overdrafts. As mentioned, after Quantitative Easing flooded the banking system with reserves, procedures in the USA were changed as the tax and loan accounts were not needed to minimize reserve impacts since banks had plenty of reserves.

Nor has the US Treasury *needed* to sell bonds to the special banks with tax and loan accounts because a system based on *dealer banks* has been created to ensure that there are always bids for bonds. The Treasury works closely with the dealer banks to ensure that it offers maturities that the banks want to buy. Typically, bond auctions are oversubscribed—meaning the lucky bidders get bonds, and those who bid too low do not. The banking system has had ample reserves to use to purchase bonds, and since bonds earn more interest than reserves do, banks usually want to exchange reserves for bonds.

It is common to state that if the government runs a balanced budget, it can use tax revenue to pay for its spending, but if it runs a deficit, it will need to borrow or print money to pay for the deficit. There is said to be a clear division between financing a balanced budget and a budget in deficit. However, as we have seen, on a daily, weekly, or even monthly basis, it is unlikely that a

budget will be “balanced,” so bonds might be sold even if the budget will be in surplus over the fiscal year. Indeed, whether the government is going to run a deficit is not known for certain until the end of the fiscal year when spending and revenue are tallied up.

What we emphasize is that day-to-day spending will follow the procedures outlined above regardless of the fiscal year outcome. What this implies is that there is no special way to “pay for” spending that depends on whether the budget will be balanced, in deficit, or in deficit (or surplus). However, the net injection of bank reserves and the net sale of new treasury bonds over the course of the year will look very different depending on the ultimate fiscal outcome.

Let’s dig a bit deeper into the “financing” question to see why there is a strong argument to be made that the budgetary outcome does not affect the *method* of financing government spending.

HOW DOES GOVERNMENT PAY FOR SPENDING?

We now face a contentious issue: how does the government *really* pay for its spending? Alternatively, we could state it this way: which comes first, taxes or spending?

Let’s begin the argument more generally: how is spending by any economic unit financed?

Innes argued that we should view all purchases as “the exchange of a commodity for a credit” and calls this the “primitive law of commerce.” (Innes 1913, p. 393)

By buying we become debtors and by selling we become creditors, and being all both buyers and sellers we are all debtors and creditors. As debtor we can compel our creditor to cancel our obligation to him by handing to him his own acknowledgement of a debt to an equivalent amount which he, in his turn, has incurred. (Innes 1913, p. 393)

We should not take this literally but as a way to think about how spending is financed: it is always based on credits and debits. It is intuitive to argue that a worker’s wages finance consumption, or that a firm’s sales revenues finance its payment of wages. The worker receives a check or a direct deposit that leads to an increase in the worker’s bank balance. What goes on behind the scenes is that the employer’s bank has debited the employer’s deposit balance and made a payment to the worker’s bank that involves a transfer from the employer’s bank reserve account at the central bank in favor of the worker’s bank. Now the worker can write a check (or use a debit card or other form of bank transfer) to buy groceries. The worker’s deposit is debited, and the grocer’s account

(probably at a different bank) is credited. That leads to a transfer of reserves from the worker's bank's reserves to the store's bank's reserves.

What if the worker tries to spend before the paycheck clears? The check submitted to the bank bounces, or an overdraft is recorded. In many nations, overdrafts are routine, although in the USA they can incur fees except for depositors who normally maintain large balances. Of course, an overdraft is equivalent to a loan (credit to the bank's assets, and liability of the worker). An alternative way to pay is with a credit card, in which case the card-issuing bank pays the store's bank and the worker's bank pays the bank that issued the card. While one's wages might be linked to the ability to pay, receipt of wages may not be required to make payments. And in all these cases, the act of paying takes the form of credits and debits to bank accounts.

In conclusion, we can ask whether the wages *paid for* the consumption spending. In our example, and more generally, the wages are not received in a form that the worker can use for spending (i.e., a check drawn on the employer's bank must be cashed first), and the worker's check is not received in a form that the store can spend. Rather, the employer's bank made a payment on behalf of the employer (received by the worker's bank which credited the worker's deposit), and the worker's bank made a payment on behalf of the worker. This is what banks do: make (and receive) payments for their customers. What if the worker had not received the wage payment? The bank could still make the payment on behalf of the worker by allowing an overdraft. We can conclude that a worker's spending is almost certainly linked to the worker's desire and ability to spend, but the worker doesn't *really* spend the wages.

What about sovereign governments—do they really need and really spend tax revenue? If we go back to the time before central banks (as we saw, first created in the late 17th century), sovereigns could directly spend their sovereign currency in the form of coins, paper money, and tally sticks. They then *redeemed* the currency in tax payments. In that case, it should be quite clear that the sovereign spent the currency first and then collected the currency in tax payments. This was the argument of Innes—while taxes create a demand and a value for the currency, they do not really finance the spending. Rather, we could say that spending “finances” taxes in the sense that spending the currency provides the means used by the taxpayer to meet the tax obligation.

We saw in Chapter 3 how the Bank of England was created as Parliament sought both to rein in the sovereign King's power and to strengthen the sovereign currency. The evolution to a system in which the central bank becomes the sovereign's bank, making and receiving all payments on behalf of the treasury, took place over a long period. Recall that England linked the currency to precious metals to constrain but also paradoxically to strengthen sovereign power. English sovereign debt was so acceptable that debt-to-GDP ratios climbed as high as those of supposedly “irresponsible” governments like Greece today.

Eventually, precious metal was no longer needed to build trust in the currency, and so the link between currency and metal was abandoned.

The developed capitalist economies now have universally adopted monetary systems based on a national sovereign currency and central banking, with flexible exchange rates. (We will see that the Euro area is somewhat of an exception.) Fiscal policy space has grown to the point that government spending commonly commands a quarter to a half of national production, with central banks making and receiving all payments on behalf of the treasury.

However, this transition has also obscured the nature of government finance. Treasuries no longer issue a significant amount of currency, and almost all payments involve two degrees of separation between the sovereign government and the public: a central bank and a private bank are involved in virtually all payments by and to the government. Yet, as discussed, overdrafts are prohibited, so modern treasuries generally need positive balances in deposit accounts at the central bank before payments will be made. It appears that tax revenue is required before the government can spend. Has the sequence been reversed so that taxes *finance* spending?

MMT has typically insisted that it has not. Let's investigate the arguments that make the case that spending should still be considered prior to taxes so that taxes really are not an operative financial constraint on government spending. To put the question more bluntly, is the government constrained by its *income* in the form of tax revenue?

First, as we've seen, wages are no longer received in a form that the worker can directly spend. Like the worker, the treasury no longer receives in payment what it spends. Under current operating procedures, the central bank credits the treasury's deposit when taxes are received and debits it when payments are made on behalf of the treasury. However, those credits and debits are not in a form that taxpayers can pay or receive. Tax receipts take the form of a debit to the taxpayer's bank deposit and a credit to the treasury's deposit at the central bank, and simultaneously a debit to the reserves of the taxpayer's bank, and a debit to the central bank's reserve liability (held by the taxpayer's bank). When the treasury spends, the central bank makes payments on its behalf, crediting a bank's reserves and simultaneously debiting the treasury's account, while the private bank credits the deposit of the recipient using its own liabilities. Just as the private bank can make payments on behalf of a worker even if no wages are received, a central bank could make payments for the treasury even if no taxes are received, so long as overdrafts are allowed. Otherwise, the treasury can obtain a deposit at the central bank by selling bonds—as the central bank will debit the reserves of the purchaser's bank, and that bank will debit the deposit of the bond buyer. (Note, the private bank itself could be the purchaser of the bond.)

This leads us to the second point. Who is more credit-worthy, the indebted worker or the indebted government? Which is more risky—an overdraft by the treasury or one by the worker? The sovereign currency sits at the top of the pyramid as the *valuta* money, and payment in *valuta* money discharges a debt. If a government makes a purchase using its own currency, the payment is final—on a flexible exchange rate there is no promise by the government to deliver anything in exchange. By contrast, if a private entity makes a payment using its own liability, it is still in debt. Even a private bank cannot pay its debts using its liabilities. Only payment using currency—usually in the form of central bank reserves—fully cancels the bank’s debt.

However, the worker can run up overdrafts while the treasury usually cannot. This prohibition is similar in purpose to that of Parliament’s constraint on the King: to build trust in the currency by creating the appearance that the modern treasury cannot spend through money creation. Yet, we know that most of the rich sovereign nations continuously spend in excess of tax revenue—regardless of the restriction. Further, governments are subject to democratic control—presidents or prime ministers are not kings and spending is constrained by budgets. Treasuries can spend up to budgeted amounts (with the notable exception of the USA, which has a debt limit)—so elected representatives determine spending amounts. It is a bit ironic that the worker gets overdrafts, but Uncle Sam cannot have them.

This is clearly a self-imposed constraint that can be removed (and sometimes is) to allow the treasury to continue to make payments even if its deposit account at the central bank has fallen below zero. Internal accounting would show the central bank holding the liability of the treasury, offset by central bank reserves created to make a payment for the treasury. (We will examine this in more detail in the next chapter.) From the perspective of the government as a whole, it is as if the government owes itself. If the treasury pays interest on that debt, it is a payment of one government entity to another—the government pays itself. Is this risky? If the government defaults, it wipes a liability and an asset off its balance sheet. Only a government accountant would notice—there would be no impact on the nongovernment sector. This is not to imply that defaulting would be a rational policy, as it would diminish trust in the government.

Third, as we have seen, the operating procedures have been developed largely to ensure the payments system operates smoothly. As Bell (2000) argued, the functional purpose of the fiscal procedures developed by the treasury and central bank is to minimize reserve effects of fiscal operations. This was more important in the USA before 2009 because the Fed did not pay interest on reserves. As we’ve seen, Treasury spending adds reserves to the banking system, so over any period that spending exceeds taxes received, net reserves are added; when taxes exceed spending, the surpluses reduce reserves. These

additions to/subtractions from reserves needed to be carefully monitored to prevent undesired impacts on interest rates, with a central bank injection of reserves used to make up any shortfall or a reserve drain used to remove excess reserves. The central bank adds reserves mostly through bond purchases; it drains reserves through bond sales. (Central bank actions to offset the daily fluctuations that would destabilize the overnight interest rate (called operating factors) will be taken up in detail in Chapter 9.)

Note that the *function* of selling bonds, whether by the treasury or the central bank, is to reduce reserves held by banks—the point made by Lerner’s second principle of functional finance. Treasury spending increases reserves, and unless the banks want more reserves, the extra reserves will be offered in the overnight market (fed funds market in the USA), putting downward pressure on the overnight interest rate. That pressure is relieved by a bond sale—and unless the Treasury happens to simultaneously sell bonds, it will be up to the central bank to engage in an open market sale of a treasury bond to drain the extra reserves.²

Because of the reserve effects of fiscal operations, it would not normally be possible for taxes to be paid before government spending because excess reserves would not be available to debit when banks make payments for taxpayers. The Treasury cannot simply withdraw taxes from the economy before spending—any transfer from tax accounts to the government’s balance sheet must be exactly offset by government provision of an equivalent amount of sovereign currency in the form of reserves through use of the Fed’s balance sheet. A “tax first, then spend” policy is not viable even in today’s economy, which is based on banking and central banking, without adopting complex operating procedures that make the sovereign currency (in the form of central bank reserves) available for tax payments.

It is best to keep separate the notion of finance versus the spending-income relationship. The finance is created simultaneously with the act of spending—it consists of a credit to a bank account and a debit to another bank account. Finance is a different category from income flows like wages, profits, and taxes, and spending flows like consumption, investment, and government spending. While a bank is more willing to make a payment for an entity that has income flows, those flows are not really the source of the payment. When a worker spends on a consumption item, there is a debit to the worker’s account and a credit to the retailer’s account. Wages are not directly involved in the payment—although spending by workers is influenced by wage income.

GOVERNMENT AS SUPPLIER OF OUTSIDE MONEY

Starting from the aggregate level, it is easy to see that only the government can be a net supplier of domestic money. Every private creation of money held

by someone is offset by the creation of an equal private liability. For example, every time a bank creates a demand deposit (bank money), it is carried on the books of the depositor as an asset, but it is a liability of the bank. As Gurley and Shaw argued, bank money is *inside* money: while it is an asset for the holder (depositor), it is offset by the bank liability and can never be a net asset of the private sector. Government creates *outside* money—money that is not the liability of any private entity.

Payments using bank money cause the bank money to shift pockets but leave it intact except in three cases: (1) if a bank loan is repaid, (2) if a bank check is presented for cash, or (3) if payments are made to the government. In all three cases, the bank money is destroyed (refluxed or redeemed), and the latter two cases also require outside, government money. When households use bank deposit money to pay taxes, a clearing drain results in fewer reserves left in bank reserve deposits at the central bank.

When the government creates sovereign currency to purchase goods and services (or assets such as gold, or to reduce liabilities such as payments owed to social security recipients), this shows up on the books of the public as a credit of sovereign currency and a debit of goods and services sold to the government (or as assets sold to the government, or as a reduction of claims on the government). This is *net money creation* because it is not offset by a private sector liability. This net money (called outside money by Gurley and Shaw) is available to pay taxes. When taxes are paid, the public reduces its outstanding tax liability, which is exactly equal to the reduction of its holding of net money (government sovereign currency, usually in the form of reserve deposits at the central bank).

In principle, then, the government first spends sovereign currency (to purchase goods, services, and assets or to provide transfer payments), or lends it (provision of reserves by the central bank). Once the government has spent or lent, then that sovereign currency is available to be transferred to the government to meet tax liabilities. As a matter of logic, the public cannot pay sovereign currency to the government to meet tax liabilities until the government has paid out (or lent) sovereign currency to the public. In a modern capitalist economy, it may appear more complex than this because most taxes are paid using checks drawn on bank deposits, rather than currency. However, this amounts to the same thing since every payment of taxes generates a reserve clearing drain. Thus, taxes cannot be paid until actual coins, notes, or reserves are injected into the economy.

If government runs a continuous surplus, it will be withdrawing the means of ultimate settlement, sovereign currency. Today, that means bank reserves will be continuously drained. Up to a point, the central bank can replenish the supply, normally by purchasing treasury bonds held by the public—but if there is a continuous surplus, all the treasury bonds will be retired. Once all

the bonds are retired, the central bank will have to buy private bonds or lend against them to put reserves into the system. While a government deficit puts safe assets into the private sector, a government surplus takes them out—or puts the private sector into debt, replacing a tax debt with another debt to the government.

In countries that appear to run continuous government budget surpluses, some combination of these central bank policies is invariably used to provide the sovereign currency the public must have for ultimate payment. In a sense, this is nothing more than an accounting gimmick—the government keeps two sets of books, the treasury's book and the central bank's book. It is only because the central bank's purchases of government bonds, private bonds, other private assets, foreign currencies, or even goods and services are not counted as part of government spending that it can appear that persistent government surpluses are possible.

The final point is to emphasize the implications of the budgetary outcome on the private sector's balance—adopting the sectoral balance approach of Wynne Godley. At the level of the economy, spending equals (and determines) income. We can divide the economy into sectors—as many as we want, but it is useful to use three: the domestic private sector (households and firms), the domestic government (all levels, although for a sovereign currency nation, the national government as currency issuer dominates), and the foreign sector (the rest of the world). While any individual sector can run a surplus (income > spending), a deficit (income < spending), or a balanced budget (income = spending), the sum across all three must be balanced. If the foreign sector is balanced (while it is a simplification, think of imports = exports as the balance—it is actually the current account, which is a bit broader) and the private sector runs a surplus (saving > investment), then the government sector must run a deficit (government spending > taxes). The government deficit generates a flow of private sector saving that can be used to accumulate net financial wealth.

The USA, the UK, and some other rich developed countries tend to run trade deficits against the rest of the world and at the same time almost always have positive domestic saving (running a private surplus). By identity, this means they must have a government deficit. For these countries, that is essentially the default position. When they do run surpluses in the government sector and achieve a private sector deficit, that position is short-lived because it is financially unsustainable for households (in particular) to spend more than their income for very long.

We conclude that:

1. It is misleading to argue that government needs to tax first and then spend the revenue. Taxes are paid in the government's currency, so it must be put into the economy before taxes are paid.
2. Current operating procedures obscure this relationship. When government directly spent the currency, it was obvious that spending provided currency to pay taxes. Today the relationship is obscured by two degrees of separation: the central bank makes and receives payments for the treasury, while private banks make and receive payments on behalf of taxpayers and recipients of government spending. Still, currency in the form of central bank reserves must be provided before taxes can be paid.
3. Government can and typically does run deficits, proving that currency (again, today mostly in the form of central bank reserves) can be supplied in excess of tax payments.
4. Government spending generates both income and net (or outside) money balances for the nongovernment sectors (households, firms, and foreigners).

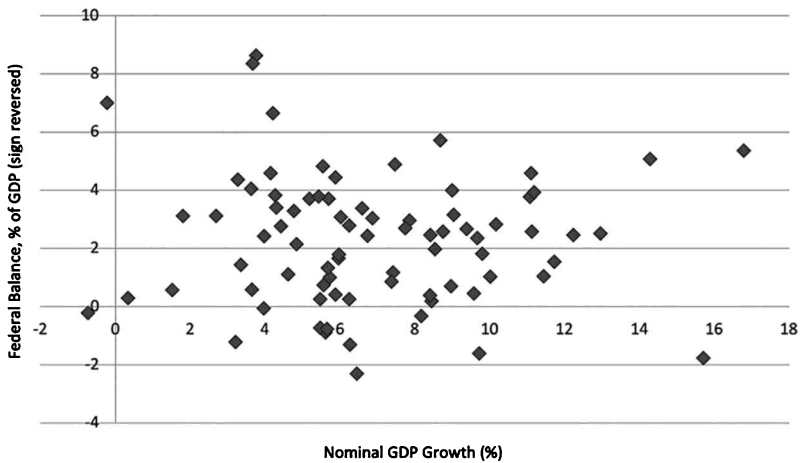
In the next section, we examine government deficits and debt in the context of economic growth.

DEFICITS, DEBT RATIOS, AND GROWTH

Orthodoxy argues that budget deficits and high government debt ratios threaten economic growth and, ultimately, are unsustainable. In large part, this is because of supposed impacts on interest rates—a topic for the next section. Empirically, then, we should expect growth to *slow* as deficit and debt ratios rise. On the other hand, the Keynesian tradition has largely recommended boosting government spending—or lowering tax rates—to stimulate economic growth. One would expect, then, that higher deficits and rising debt ratios should be associated with *faster* growth—and then the danger is inflation caused by excess demand. In this section, we will see that the relationship among deficits, debt ratios, and growth is more complex.

Figure 8.1 plots the USA Federal government deficit ratio (deficit scaled by GDP) against the growth rate of GDP.

There is no obvious pattern. Why is that? Because there are two ways to produce a budget deficit—the “good way” and the “ugly way.” Government could embark on a discretionary expansion of spending—say to finance a Green New Deal—resulting in a bigger deficit ratio. We could call that a “good” deficit. But a deep recession with slow growth could be associated with a rising deficit ratio because tax revenues fall—an “ugly” deficit. In practice, the budgetary outcome is largely nondiscretionary. Some government spending is



Note: Data excludes outliers; graph by Eric Lin, Bard College.

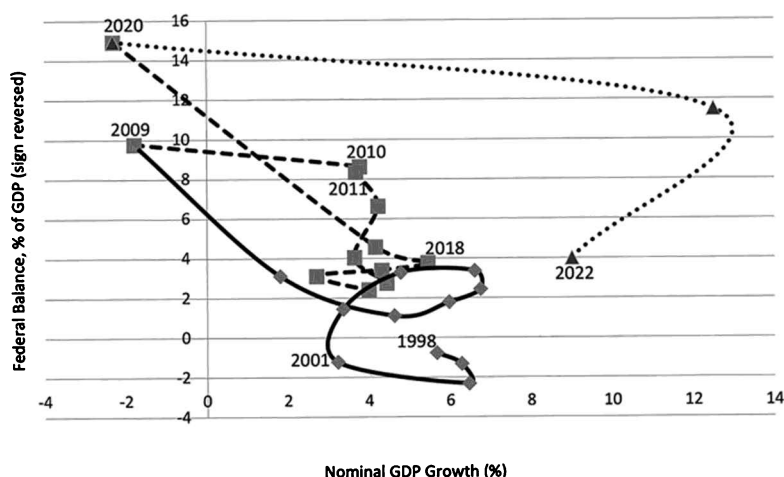
Source: BEA and authors' calculations for GDP growth rate; Office of Budget Management for Federal balance (retrieved from FRED).

Figure 8.1 Growth and the federal budget balance, 1930–2022

automatically countercyclical: it goes up in bad times and down in good times. For example, unemployment compensation and welfare spending move countercyclically (falling in expansion and rising in recession). On the other hand, most government taxes are procyclical: revenues rise in good times as income grows, and tax revenues fall in bad times as unemployment increases and sales and income fall. Once we consider these automatic and cyclical movements, we recognize that the correlation between growth and debt ratios could go either way. For that reason, we need to trace the movement of the relationship over time. It turns out that there is a very clear pattern through time.

While we will look only at USA data, a similar exercise for other countries finds a very similar pattern.³ Figure 8.2 shows a graph of the data, by year, over the past couple of decades.

Several things stand out. Under President Clinton, the Federal budget went into surplus (the deficit ratio turned negative) in the late 1990s as the economy grew relatively rapidly. The surplus was celebrated as a great achievement, but it began sucking income out of the private sector (which was already running a large deficit). The headwinds increased with the collapse of the Dot.Com bubble in 1999, slowing GDP growth through 2001. That slow growth reduced tax revenue, returning the government's balance to a budget deficit. The injection



Note: 2009, 2020, 2021 are outliers.

Source: Graph by Eric Lin, Bard College; GDP data from BEA and budget data from Office for Budget Management for the Federal balance (retrieved from FRED).

Figure 8.2 Growth and the Federal budget balance, 1998–2009, 2009–2020, 2020–2022

provided by the government deficit helped growth to pick up in the early 2000s. At the same time, housing, commodities, and stock market bubbles added to the boost—restoring economic growth and boosting tax revenue, eventually reducing the deficit ratio. By 2007, the growth was fast enough to move the budget close to balance again. That meant the automatic stabilizer was pulling demand out of the economy. This slowed the economy, and it crashed into the GFC, with growth going negative in 2009 and the budget deficit rising sharply as tax revenue fell sharply.

After the budget deficit reached 10 percent of GDP (a postwar record), the economy finally began to recover in 2010. While there was some discretionary fiscal stimulus (approximately \$800 billion spread over two years), most of the deficit was created by automatic stabilizers (tax revenues fell sharply while social spending rose), which stimulated renewed growth. Higher growth increased tax revenue, so the deficit fell sharply over the next few years (to just 2.5 percent of GDP by mid-decade) as growth remained steady.

Then the COVID pandemic hit the economy hard in 2020. Lockdowns reduced income and spending, causing an unprecedented increase in

unemployment as growth turned sharply negative. Tax revenues fell, and COVID relief packages helped to boost the deficit to another, even higher, postwar record at 15 percent of GDP. This time, it was due both to the automatic stabilizer and also to the discretionary spending increase (a total of \$5 trillion spent on COVID relief). Rapid growth caused the deficit ratio to fall sharply to about 4 percent by 2022.

While this period was indeed unusual—with the two deep recessions—a similar pattern (although with smaller swings) is found in the USA for earlier years and for a large number of countries over the postwar period. Generally, deficits increase in bad times, helping the economy turn around to grow more quickly. That boosts tax revenue and reduces the deficit ratio until the leakage of income to taxes eventually slows growth sufficiently that the deficit rises again. This “loop-di-loop” repeats itself over and over.

Mainstream economists, as well as policymakers, worry that without taking action to curb deficits and debt, they might grow faster than GDP, raising debts and deficits to a high level. And if deficits and debt really do hinder economic growth, we will find ourselves with rising deficit *ratios* and a GDP *growth rate* that continues to fall. If we carry those trends for decades, centuries, and even millennia, the debt ratio goes to “infinity and beyond.” However, the evidence indicates that outcome is quite unlikely because deficits do act in a Keynesian manner: in bad times, they put a floor to growth and help it to recover, and in good times, faster growth automatically reduces deficits and moves the budget toward surplus.

This does not mean that we should never take discretionary action. The automatic stabilizers may not be sufficiently powerful to prevent inflation when the economy grows too fast, or high unemployment when it grows too slowly. Furthermore, what the USA (and the world, more generally) has seen over much of the period since 1980 is “jobless growth”: even with reasonably high growth rates, not enough jobs are created. To some extent, that is due to higher labor productivity, and in part due to “robotization”—and AI (artificial intelligence) could accelerate that trend. Furthermore, there is no guarantee that the automatic stabilizers are directed where they might do the most good. Rising tax revenue might result from rising tax rates on wages rather than on interest, profits, and rents that go to high income households; rising government spending might take the form of higher interest payments by government to high wealth households as budget deficits generate more government debt. The automatic stabilizers might increase the inequality of income and wealth over the cycle.

Since the government is the sole supplier of sovereign currency, and since sovereign currency is essentially a public resource that is potentially unlimited in supply, it makes little sense to restrict government spending and the supply of sovereign currency to the extent that this causes unemployment, unless

unemployment serves a useful purpose. We will return to this issue later; however, it is important to note that until full employment of resources is reached, there is no *real* constraint on government spending. Beyond that point, real constraints intervene so that additional spending will run up against the inflation barrier. In Chapter 11, we will examine a policy proposal that would allow the pursuit of true full employment without generating inflation pressure.

IS GOVERNMENT SPENDING CONSTRAINED BY BORROWING THROUGH BONDS SALES?

Some economists and policymakers worry that so-called bond vigilantes might refuse to *lend* to government by purchasing treasury bonds if they believe the government's debt ratio is too high. We argue that government spending is not constrained by the quantity of bonds that markets are willing to purchase; rather, the functional purpose of bond sales is to provide an interest-earning alternative to cash and excess reserves. Understanding how the government really spends, and thus, why government issues bonds, can help us understand why we need not fear a strike by bond vigilantes.

While the central bank is often prohibited from buying government bonds in the primary market (new issues), it can buy them in the secondary market. When the central bank buys bonds in the secondary market, it credits bank reserves. Since banks would prefer to hold higher-earning assets, they will seek to exchange reserves for the bonds. The central bank can thus support the market for government bonds by purchasing them in the secondary market—ensuring bank willingness to buy them. The central bank can always make more space for newly issued bonds through open market purchases (in the secondary market), putting extra reserves into banks. And, in fact, in the USA, dealer banks must place bids for new issues—so the only question is what price they will bid. Adding more excess reserves to the banking system will hold down the interest rate required of new bond issues (equivalently, increase the price bids of the so-called vigilantes).

Thus, government spending is constrained only by its budget (in the USA, the budget is formulated by Congress) and private sector willingness to provide goods, services, or assets to the government in exchange for government money: anything that is for purchase in terms of the domestic currency can be had through the creation of sovereign currency.

Governments sometimes believe that they must sell bonds in international markets because domestic markets are already saturated, and any further domestic sales would require higher interest rates. While it could be true that at a higher interest rate, the government might be able to sell more bonds, this is not an indication that the government is *forced* to pay higher rates to finance its deficit. Once domestic households and banks are content with their holdings

of government bonds, then the government need not drain any more reserves from the system—or, equivalently, it need not sell any more bonds. In practice, if a deficit is associated with rising interest rates it is because the central bank is raising its interest rate target. All else equal, deficits place *downward* pressure on overnight rates—because they add sovereign currency in the form of bank reserves—but bond sales relieve the pressure. Indeed, that is the functional purpose of selling them.

What if the government does sell bonds to foreigners? So long as these bonds are denominated in the domestic fiat currency, they do not entail any risks that domestically held bonds do not hold—and they serve the same purpose, which is to provide an interest-earning alternative to non-interest-earning sovereign currency (in the form of central bank reserves). When China accumulates reserves at the Fed (due to net exports to the USA), the Chinese central bank asks the Fed to move the reserves into US Treasuries to earn interest. (The Fed keeps the accounts for the Treasury—this is essentially just like moving funds from your checking account to a savings account at your bank to earn a higher interest rate.)

Concern is often expressed that foreign-held bonds commit a country to paying interest to foreigners, which is said to “burden” domestic taxpayers (see also the discussion of burden of the debt below). However, interest can always be paid through the creation of sovereign currency (again, in the form of central bank reserves; if the interest is earned by an entity that is not a bank, the bank’s reserves are increased and the bank credits the deposit account of the bondholder)—as the central bank makes payments on behalf of the treasury—just as any other government spending is financed through creation of sovereign currency.

Some worry that interest payments will lead to larger budget deficits, meaning more bond sales in the future. The debt will continue to grow and supposedly burden future generations even more! But future bond sales will operate functionally to drain excess reserves, just as current bond sales drain reserves; so, this entails no “burden.” And there is no possibility that the government might find itself in a crisis because it is unable to sell or roll over bonds—since the functional purpose of bond sales is to offer interest-earning alternatives to sovereign currency. As Lerner argued, bonds should only be sold when the public has more (non-interest-earning) currency (cash and bank reserves) than desired.

Whether or not the USA *should* pay interest to foreigners is a different question—the USA could simply refuse to sell the bonds, forcing foreigners holding reserves at the Fed to forego interest. This *might* reduce foreign demand for US dollar reserves, although that is not certain. Dollars are held abroad for a variety of reasons, including to protect domestic currencies. Many nations look to the USA as a market for exports. In both cases, those countries would

probably be willing to forego interest earnings and simply hold reserves at the Fed. This argument might not apply to other nations that do not issue the international reserve currency or whose economies are so small that they are not seen as good markets for exports. In that case, paying interest on bonds could maintain a higher demand for the currency, and this might be a policy such nations would want to pursue. (See below.)

Sometimes governments believe that the market forces them to issue *foreign-currency-denominated bonds*. There is one case in which this would be true—when the government wishes to purchase goods and services that are not for sale in terms of the domestic sovereign currency. In this case, the government cannot issue the currency desired by sellers. Nor can it dictate the price it will pay in terms of its own sovereign currency. In this case, the government must exchange its currency for foreign currency, sell something abroad, or sell a bond in the foreign currency to obtain the currency required by sellers of the goods and services it would like to obtain.

Unlike the case of domestic-currency-denominated bonds, this can burden an economy because interest payments in foreign currency cannot be financed through creation of sovereign currency. The government will need to obtain foreign currency to make the payments. The government may have little influence over the foreign-currency price at which the bonds will sell (which reflects the interest rate paid in an inverse manner: a lower price implies a higher interest rate). Additionally, it will have to obtain additional foreign currency in the future to service the debt. In some situations, markets might fear that a government will not be able to obtain the currency needed to service the debt, causing a rational run out of these bonds. As a result, the government may feel forced to impose austerity on its population to maintain a trade surplus and obtain the needed foreign currency. In this case, the austerity can be at least partially blamed on “market discipline.” However, it must be recognized that this is because the government desired goods and services that were not for sale in the domestic currency. In all other cases, the government is not subject to “market discipline”, and any austerity and hardship is self-inflicted.⁴

Since domestic currency-denominated bond sales are functionally an interest rate maintenance operation, the government decides the interest rate it will pay—or, alternatively, the price of government bonds—through its monetary policy. Thus, sovereign governments are not subject to market discipline regarding either the quantity of bonds sold to private holders or the price at which they will be sold, so long as the bonds are issued in the domestic currency. When there are excess reserves, the market will demand bonds at any interest rate above zero (or the support rate paid by the central bank), as the alternative is to hold reserves. There may be very good reasons for maintaining a significantly positive interest rate on government bonds (denominated in the domestic currency), but it is never necessary to do so merely because the

market would like a high interest rate. A high rate on government bonds is evidence that the government (treasury and/or the central bank) has chosen a high interest rate—it is not a function of market forces of supply and demand. The government can always have a lower interest rate merely by reducing the interest rate target—to zero at the extreme. (More on this in the chapter that follows.)

Most of the time, governments do not appear to understand this, but there are exceptions. For example, during WWII, the US government was able to “borrow” to finance massive deficits (up to 25 percent of GDP; only the COVID-era deficits reached as high as 15 percent in the postwar period) at a short-term interest rate of $\frac{3}{8}$ of 1 percent. At the time, the government did understand that lack of finance was not a problem. Bond sales were designed to encourage net saving by the public to ensure that goods and services would be released from consumption for the war effort (documents from the time prove this was understood by the government, and advertisements for savings bonds emphasized that the purpose was to postpone consumption to prevent inflation). The public viewed the bonds as a patriotic (and interest-earning) means of accumulating financial wealth that would be used in the postwar period to support higher levels of consumption. Clearly, the government did understand that the price of bonds (and, therefore, the interest rate paid by government) was not determined by markets; it was set by government policy. The government also understood that many of the prices it paid for goods and services were not market-determined—but, rather, were administratively set by the government. Similarly, the government determined the price of labor—directly and indirectly—so that even as the economy operated well beyond full employment, prices of labor and of output purchased by the government were relatively stable. Most of the lessons that could have been learned from WWII finances were not learned, or were forgotten.

FUNCTIONAL FINANCE AND THE BURDEN OF THE DEBT

In his *State of the Union* message of January 7, 1960, President Eisenhower made a statement that could have been made by President Jackson a hundred years earlier (and perhaps three decades later by President Clinton): “Personally, I do not feel that any amount can be properly called a surplus as long as the nation is in debt. I prefer to think of such an item as a reduction on our children’s inherited mortgage.” (President Eisenhower, quoted in Bowen, Davis, and Kopf 1962, p. 701) Economists attempted to defend Eisenhower’s belief that government debt represents a burden on future generations, arguing that future generations would have to reduce consumption to pay the taxes required to retire the outstanding debt. This was a pointed response to those

who adopted the functional finance approach, arguing that any real burden of government deficits must be borne at the time the “borrowed” funds are spent.

For example, Bowen, et. al. argued that the functional finance approach ignores the eventual burden on the later generation (dubbed Generation II) when

the government decides to retire the debt by levying a general tax in excess of current government spending and using the surplus to buy up the bonds that are now held by members of Generation II. The inevitable outcome of this decision is a reduction in the lifetime consumption of Generation II. (Bowen, et. al. 1962, p. 701)

They claimed that while it is true that the taxes raised from Generation II are merely redistributed to bondholders in Generation II, there is still a net lifetime reduction in the consumption of Generation II since these bondholders had to reduce consumption when they purchased bonds from Generation I. Such views still dominate the discussion of the long-term effects of budget deficits—they impoverish future generations that must pay off the debt.

In a masterful response, Lerner (1961) addressed many of the errors of the Bowen, et al., exposition:

The “red herring” nature of having the Lowells lend the money now (so that we can call them the present generation) and having the Thomases pay the taxes in the future (so that they can be called the future generation) jumps to the eye if we note that the shifting of the real burden of the project from the Lowells to the Thomases (or indeed of any other burden) could take place just as well at the time of the project (or at any other time) by simply taxing the Thomases instead of the Lowells. (Lerner 1961, p. 139)

The key to the Bowen, et al. result is a series of misunderstandings and strange assumptions. For example, they presume that when Generation I reduces consumption to buy bonds that result from deficit spending, there is no flow of income that results from the government spending (money that is supposed to have been raised from the bond sales mysteriously disappears). But deficit spending must (all else equal) create the same amount of aggregate saving (initially in the form of Treasury checks). Bond sales then merely allow households (and banks) to hold bonds rather than money. This does not mean that there is no burden associated with the government spending—presumably, real goods and services are moved from the private sector to the public sector—but the *real* burden is borne immediately and results whether or not the government sells bonds.

Further, the supposed burden imposed in the future by taxing Generation II is due to the government’s decision to run a surplus (taxing more than it spends so that it can retire bonds). No one who adopts the functional finance approach

would deny that a government surplus (all else equal) reduces private sector income and net saving. When the government then retires bonds, it substitutes non-interest-earning government liabilities for interest-earning liabilities; this simply restores the sovereign currency removed when the government taxed more than it spent. The same result could be obtained if the government were to forgo the budget surplus but were to confiscate (or default on) all outstanding government bonds. Yes, there is a real burden involved, but it has nothing to do with the method of financing government spending chosen. No government that understood the principles of functional finance would confiscate or default on outstanding sovereign government debt (nor would it believe it must increase taxes to retire outstanding debt). The power to tax is the power to destroy (net private sector wealth, that is).

THE EMU AND GOVERNMENT DEFICITS

In the 1990s, the European Union (EU) developed and implemented an Economic and Monetary Union (EMU) that created a European Central Bank (ECB) responsible for the European System of Central Banks (ESCB), with existing national central banks becoming operating arms of the ESCB (see Arestis and Sawyer, 1998). Members would give up their own currencies, to be replaced by a single currency—the Euro. Under the EMU, monetary policy would be divorced from fiscal policy, with a great degree of monetary policy independence to focus on the primary objective of price stability. Fiscal policy would be tightly constrained by criteria dictating maximum deficit-to-GDP and debt-to-GDP ratios (3 percent and 60 percent, respectively).

Most importantly, as Goodhart (1997) recognized at the time, this would be the world's first modern experiment on a wide scale that would break the link between a government and its currency. The EMU would have a single currency and essentially a single monetary policy set by the ECB but would (in theory) allow each country to operate an independent fiscal (spending and taxing) policy (within the deficit and debt constraints). Most importantly, it was envisioned that this arrangement would forever eliminate the possibility that a member government could use monetary policy to bail out errant fiscal policy, for example, by having the central bank “print money” to finance excessive deficits. Fiscal policy became subject to the Maastricht criteria (with penalties imposed for excessive deficits) and to the dictates of the market (since treasuries would be limited in their ability to turn to central banks to purchase the government bonds that private markets might not want).

It was sometimes claimed that the EMU simply followed the example of American integration, with a single currency used throughout the United States (thereby reducing transactions costs and contributing to integration of the member states). Goodhart (1997) and Godley (1997) rightly questioned

the analogy. In the United States, as we have discussed, what are normally called monetary and fiscal policy are in fact closely intertwined—requiring close cooperation between the Fed and Treasury. Government spending is financed through issue of currency, taxes generate demand for the currency that results in sales to government, bond sales substitute bonds for central bank reserves, and central bank operations determine interest rates and defensively add or subtract reserves. The Federal government is not constrained by Maastricht-like criteria, nor by financial markets. The only financial constraint faced is the budgeting process adopted by Congress—and to some extent by Congress's self-imposed debt limit (that until recently was usually raised without too much fuss).

The relationship of member countries to the EMU was more like the relationship of the treasuries of member states of the United States to the Fed than it is of the US Treasury to the Fed. In the US, states have no power to create currency, and states really do have to borrow (sell bonds to markets) to spend in excess of tax receipts. All but two US states have balanced budget requirements. Purchasers of state bonds do worry about the creditworthiness of states, and the ability of states to run deficits depends at least in part on market perceptions of riskiness.

In addition, US states distinguish between the current account (consumption-like spending such as paying salaries to teachers and police) and the capital account (investment-like spending on infrastructure such as bridges and football stadiums). The capital spending is then expensed over a number of years, so only a small fraction of the cost shows up in one year's current account. While the current account must be balanced each year, the state (and local governments) can borrow to finance capital spending. Of course, debt must be serviced, and interest payments are included in the current account. If a state does end up with a deficit on its current account, it is supposed to reduce spending or raise taxes in the subsequent fiscal year.

The US federal government does not use a capital account; its investment-type spending is expensed in the fiscal year and shows up in the current account. As the currency issuer, it does not really need to borrow and can make all payments as they come due, so it (arguably) makes sense for the national government to dispense with a capital account. But the new Euro nations were to become more like US states. At the time of the launch of the Euro, no state in the USA had a government debt ratio as high as 20 percent of the state's own GDP—while all but one of the new EMU members had higher ratios—even up to a reported ratio of 100 percent or more in the cases of Italy and Greece. (It was later found that the debt ratios were even higher than reported as they were hiding debt.) As Goodhart (1997) pointed out, the US state with the very worst credit rating in 1994 was Louisiana, with a ratio of debt service to revenues of 11.5 percent. In contrast, Germany had a ratio of 17.8 percent, while Italy

reported a ratio of 50 percent—and presumably no one operating in private markets would believe that Louisiana was a better credit risk than Germany or even Italy. The reason, of course, is that markets recognize that the treasuries of sovereign nations do have the ability to issue currency and to service debt, while the individual US states do not. However, once the EMU members gave up their currencies, default risk emerged.

In some respects, US states are even safer than EMU members were under the new arrangement: US states are backed by the spending power of Uncle Sam. If a crisis should hit the nation (which it did with the Global Financial Crisis, and later with the COVID recession), Washington would provide social assistance, ramp up federal spending, and bail out the financial system (including state-chartered and regulated banks). It was not clear that the individual members of the EMU would be so fortunate—indeed, the Treaty has a “no bail-out clause” and member nations are prohibited from assisting one another.

As designed, the EMU would have a central bank (the ECB) but no self-financed fiscal branch. The European Union Parliament was provided with a very small budget (about 1 percent of EU GDP), and that budget is contributed by member nations. So, for the region as a whole, this is not net spending—unlike the “dollar area” in which the Federal government could run up a deficit of 10 percent of national GDP, spreading its spending around the country in the aftermath of the GFC.

The setup of the EMU, then, would be much like a US that operated with a Fed, but with only individual state treasuries. At the time, MMT critics as well as Wynne Godley argued that it would be as if each EMU member country were attempting to operate fiscal policy in a foreign currency, so deficits would require borrowing in that foreign currency according to the dictates of private markets:

If a government does not have its own central bank on which it can draw cheques freely, its expenditure can be financed only by borrowing in the open market in competition with businesses, and this may prove excessively expensive or even impossible, particularly under “conditions of extreme emergency” ... The danger, then, is that the budgetary restraint to which governments are individually committed will impart a disinflationary bias that locks Europe as a whole into a depression it is powerless to lift. (Godley, 1997, p. 2)

MMT proponents, as well as Goodhart and Godley (and a few others), all warned that the setup of the EMU would not be able to support stable economic growth because adequate countercyclical spending in a downturn would be difficult. The nations would not be able to deal with a deep recession or a financial crisis because their deficit and debt ratios would grow beyond the limits and beyond what markets would accept. We warned that interest rates on government debt would explode, raising deficit and debt ratios further due

to skyrocketing interest payments. At that point, credit markets would shut governments out of financial markets.

The only solution in such a case would be either to leave the monetary union and restore a domestic sovereign currency, or for the ECB to come to the rescue. As the GFC spread to the Euro area, it looked like Greece and perhaps other “periphery” countries (like Portugal, Greece, and Italy) would either choose to leave or be forced to do so. Even though an ECB rescue was explicitly prohibited by the agreements, that is exactly what happened in 2012 when the head of the ECB, Mario Draghi, announced that he would “do whatever it takes” to back up the sovereign debt of individual members. That finally resolved the crisis that had threatened the survival of the EMU. The backstop was reinforced during the COVID-induced recession with the ECB supporting the deficits of members and keeping interest low on their debt. The EMU system was saved. However, the divorce of fiscal policy from national sovereign currencies remains a drag on economic growth in most of the Euro area.

Before the GFC, Kregel (1998) had proposed a solution to the likely chronic deflationary forces that would be created by the EMU. He would have the ECB offer to provide enough Euro funding to “hire all workers who were willing and able to work, but could not find employment in the private sector, at a basic wage.” (Kregel, 1998, p. 10) This is a Job Guarantee program, outlined below in Chapter 11, with the ECB taking over the fiscal policy role. Whenever private markets in the individual member countries were depressed, the ECB would provide the funding necessary to return to full employment; when private markets recovered, the ECB’s spending would fall. In this case, rather than having the individual member countries follow the principles of functional finance, it would be the ECB’s role to fund the spending required to reach full employment without regard to the principles of so-called “sound finance.”

While it is unusual for central banks to provide this kind of fiscal support, there did not appear to be another viable option. Individual nations could not let their spending expand as necessary because they would be punished in bond markets and also due to the Maastricht criteria. However, now that the ECB is standing behind individual country debt, if the deficit and debt ratio criteria were relaxed, each could afford such a program without worrying about the interest rates on their bonds rising. The current backstop might be adequate to allow members to relax fiscal constraints; however, most are apparently unwilling to do so. They typically operate well below full employment levels, probably to avoid larger deficits and debts. In other words, they cannot be certain that the ECB will rescue them, so they act as if they remain subject to the rules.

The next severe downturn will again test the ECB's willingness to "do whatever it takes," but it is likely that it will—again—come to the rescue of member nations.

NOTES

1. The deficits still might be inflationary either due to bottlenecks that lead to supply constraints and cost-push inflation, or due to excessive aggregate demand (causing demand-pull inflation) if the economy operates beyond full capacity.
2. Since 2009, however, the Fed has paid interest on reserves, so excess reserves push the market rate down to the rate paid by the Fed. On the other hand, if the Treasury runs a surplus, that can leave banks short of reserves, causing them to bid the overnight rate above target. This is relieved by a central bank purchase of bonds.
3. For more discussion of the results found for a number of countries, see "The Relation Between Budget Deficits and Growth: Complicated but Clear", by L. Randall Wray and Eric Lin, a Working Paper of the Levy Economics Institute of Bard College, <https://www.levyinstitute.org/publications/the-relation-between-budget-deficits-and-growth-complicated-but-clear>.
4. This does not imply, however, that there can never be a run on a country's currency, including a run on its government debt—not due to fear of default by the government, but due to fear of currency depreciation. It is possible that a large government deficit could trigger fears of depreciation. So, while we deny that a deficit by itself can generate a rational fear of default on domestic-currency-denominated government debt, we do recognize that deficits can impact expectations concerning the international value of a currency.

9. The balance sheet approach to money and finance

INTRODUCTION

In this chapter, we take a balance sheet approach, using simple transaction accounts (T-accounts) to summarize and clarify the endogenous approach to money and finance. We will see how banks advance credit to cover spending and then recapture deposits and use central bank reserve accounts for clearing. Thus, banks need not passively wait for deposits that create excess reserves before they make loans—as in the old textbook view. We also show that credit can finance investment expenditures that generate sufficient aggregate profits (as in the Kalecki equation) to enable the debt to be retired. This chapter also clarifies what velocity means in a model of endogenously created money and shows the role that government deficits play in allowing economic growth. We will discuss the accommodative behavior of central banks, contrasting that to the old deposit multiplier story. Finally, we provide a detailed look at the balance sheets involved in government spending—following the USA procedures discussed in Chapter 8.

A SIMPLE MODEL

In this section, simple balance sheets are used to demonstrate some of the propositions concerning an endogenous money supply. Minsky (1986) characterized money creation as simply a balance sheet operation. Following his suggestion, we show how spending can be financed through the expansion of banking system balance sheets.¹

Assume that the public holds its wealth in one of five forms: goods and services (GS), demand deposits (DD), government bonds (GB), time deposits (TD), and long-term private securities (PB). The central bank requires that commercial banks hold required reserves of ten percent against demand deposits, but no reserves against time deposits. In this first, simple model, there is no currency, but it will be incorporated within the expanded model below.

We assume the public sells labor services to firms, receiving a demand deposit. When a transaction occurs, we only look at the end of the period to

determine the quantity of demand deposits the public has accumulated, which equals the value of the services sold. This approach is taken so that we will not need to present income statements in addition to balance sheets. Thus, for each sector, we present the end-of-period balance sheet that allows time for income and spending flows to accumulate to stocks.²

Assume that commercial banks hold reserves (RR), short-term and long-term private securities (commercial paper, CP, and private bonds, PB), and government debt (GB) as assets. Assume that banks do not hold excess reserves, and that all reserves are held at the central bank (Fed). In the simple model, net worth will be assumed to be zero initially for all economic units except the public.³ Assume that five sectors exist: treasury, central bank, commercial banks, public, and firms (of which some are consumption sector firms and others are investment sector firms). The public will consist of workers whose marginal propensity to consume out of income is one, while firms consist of capitalists, whose marginal propensity to consume out of income is zero. These restrictions can be dropped but are retained here to simplify the analysis. We initially assume that capital asset holdings are zero but will later introduce investment and capital accumulation.

For the public, the end-of-period demand deposits held represent the accumulated income (wages) received for the sale of a stream of labor services. The public holds savings in the form of demand deposits, time deposits, and private bonds.⁴ We treat labor as the only resource required, supplied on demand. We do not debit the worker's balance sheet for supplying labor services, so the receipt of wages adds to net worth. (We implicitly reject the utility theory that treats work as disutility equal to the money value of the wage.) Workers produce both consumption goods and investment goods. When they spend demand deposits on consumption goods (in a section below), we again focus on the balance sheet statement for the end of the period. Thus, the workers hold consumption goods rather than demand deposits. Since the total wages paid to produce both consumption goods and investment goods gets spent on consumption goods, profits are generated in the consumption sector. The profits are equal to spending on wages in the investment sector. This is consistent with Kalecki's model, and also with Marx's approach in which surplus value is generated in production (see Chapter 10). With only two sectors, only one can be profitable—there are no profits in the investment sector. For the purposes of these simple models, this approach is adequate.

Balance sheets for each sector can be formulated as follows, in Figure 9.1 and 9.2.

Firms issue securities (PB, CP) to obtain financial assets (DD, TD) or to purchase labor services from the public (LS). The treasury issues securities (GB), purchases goods and services from firms, holds demand deposits at the Fed (DDg), and imposes taxes on the public (T or Tax). Banks purchase

Commerical Bank		Federal Reserve	
Asset	Liability	Asset	Liability
RR	DD	GB	RR
GB	TD		DDg
CP			
Treasury		Firms	
Asset	Liability	Asset	Liability
DDg	GB	DD	PB
GS		TD	CP
T			
Public			
Asset	Liability		
DD	NW		
TD	Tax		
GB			
PB			
LS			

Figure 9.1 Balance sheets for each sector

financial assets issued by firms (CP, PB) or by the government (GB). They issue liabilities (TD, DD) to finance the purchase of these assets. Firms use the labor services purchased from the public in the production process to generate a flow of income over time. If all goes well, this income flow will be sufficient to allow the retirement of debt issued to undertake production and to generate profit flows that can be accumulated as net worth.

In Figure 9.2 we show what happens when firms decide to purchase labor services (LS) from the public by issuing short-term debt to commercial banks (see Step 1). This increases the public’s net worth (NW), held in the form of demand deposits (DD). Note that we do not show the sale of LS as a negative

Step 1: Bank Issues Loan

Commerical Bank		Firm		Public	
Asset	Liability	Asset	Liability	Asset	Liability
+CP	+DD	+LS	+CP	+DD	+NW

Step 2: Bank Issues Time Deposit to Return to Equilibrium

Commerical Bank		Firm		Public	
Asset	Liability	Asset	Liability	Asset	Liability
CP	-DD +TD	GS	CP	-DD +TD	NW

Step 3: Firm Issues Bond to Retire Loan

Commerical Bank		Firm		Public	
Asset	Liability	Asset	Liability	Asset	Liability
-CP	-TD	GS	-CP +PB	-TD +PB	NW

Figure 9.2 Bank lending finances production, firm issues bond to retire loan

entry on the public's balance sheet. The firm will convert the LS to a good or service (GS) for sale (to be shown later).

Because banks hold required reserves equal to 10 percent of demand deposits, this cannot be a position of equilibrium.⁵ However, banks can return to equilibrium by inducing the public to exchange its holding of demand deposits for interest-earning time deposits (on which there is no reserve requirement). (See Step 2.) This may require an increase in the interest rate on time deposits to induce the change in the portfolio of the public. (Alternatively, the central bank could lend reserves.)

Generally, firms prefer to issue longer-term liabilities if the assets they hold are long-term. Thus, the firm in this example would like to sell a long-term security (private bond, PB) so that it can retire the short-term security it originally issued (CP). (See Step 3.) The public is induced to draw down its time deposit to purchase the long-term securities issued by the firm. The firm must be willing to pay a higher interest rate than the banking system pays on time deposits if the public prefers time deposits and/or incurs a penalty when it shifts funds to private securities. (We have ignored the intermediate step, in which the public first exchanges time deposits for demand deposits and uses these to purchase the securities. If this process is not instantaneous, the banking system will temporarily experience a rise in required reserves. If reserve requirements are calculated based on average demand deposits held, then the banking system will be forced to increase its reserves.)

These simple balance sheets have shown that firms are able to expand expenditures by first issuing short-term liabilities. If the banking system is willing to purchase them, it can issue demand deposits to the firms, which are transferred to the public in exchange for services rendered. Equilibrium for the banking system is restored when the public is induced to exchange demand deposits for time deposits. Firms retire short-term debt by issuing long-term debt (bonds). In the example above, the public bought the long-term debt, but the results would not be significantly different if it were sold directly to banks. Indeed, it is sometimes said that banks take up the long-term debt that households do not want to hold. (As discussed in Chapter 7, Paul Davidson's "m" is the marginal propensity to purchase bonds out of saving flows, and if it is less than 1, banks must take up the extra.)

This example shows that it is possible for credit creation to be the basis for expenditure. No exogenously supplied money was needed in these examples to allow expenditures to proceed and balance sheets to expand. However, we ignored required reserves, to which we now turn.

RESERVE REQUIREMENTS

Reserves are required on demand deposits that are temporarily created before the banking system can induce the public to exchange demand deposits for time deposits. In the USA, reserves have been calculated based on average demand deposits (and short-term business saving deposits) held over the relevant period. The quantity of reserves required would thus be a function of: 1) the reserve requirement ratio; 2) the quantity of demand deposits created; and 3) the average length of time the demand deposits exist. The faster banks can induce the public to give up demand deposits, the lower the quantity of reserves required.

In the example above, required reserves will increase due to the rise in demand deposits temporarily issued. As has been argued in earlier chapters, banks can meet such reserve requirements by borrowing at the discount window, by purchasing excess reserves in the Fed funds market, by borrowing from overseas banks in the Eurodollar market, or through various reserve-economizing activities.

Simple Numerical Example

Retaining all the assumptions above, assume the relevant period over which reserve requirements are calculated is four (five-day) weeks. Assume that a firm wishes to make a \$1000 purchase of labor services from the public, initially financed as above by an issue of short-term liabilities of \$1000 to the banking system. This is shown in Figure 9.3, Step 1.

Assume the firm holds a demand deposit for one day before it purchases labor services, and that it takes four days for the banks to induce the public to substitute time deposits for demand deposits. Assuming a 20-day period,

Step 1: Bank Makes a Loan and Firm Hires Labor

Commerical Bank		Firm		Public	
Asset	Liability	Asset	Liability	Asset	Liability
+1000 CP	+1000 DD	+1000 LS	+1000 CP	+1000 DD	+1000 NW

Step 2: Bank Borrows Reserves to Meet Required Reserve Ratio and Firm Sells Bond to Retire Loan

Commerical Bank		Firm		Public	
Asset	Liability	Asset	Liability	Asset	Liability
+30 RR	+30 BR	+1000 GS	-1000 CP	-1000 TD	1000 NW
-1000 CP	-1000 TD		+1000 PB	+1000 PB	

Figure 9.3 Commercial bank borrows reserves to meet requirements

required reserves would rise by $(0.10)(0.25)(1000) = \$25$. If firms sell long-term securities (PB) to the public to retire the short-term debt (CP), and this process takes an additional day (the public must first convert TD to DD, then purchase the PB), then required reserves will rise by $(0.10)(0.30)(1000) = \$30$. Following those transactions, the balance sheets will look as shown in Step 2.

Note that BR stands for reserves the commercial bank has borrowed to meet the reserves required on the basis of the demand deposits temporarily issued. Interest payments have been ignored, and the balance sheet of the bank excludes any increase in net worth. The bank would not have undertaken this operation without receiving interest, since it must pay interest on the reserves borrowed to meet required reserves. Required reserves are similar to a tax: they increase the cost of issuing liabilities. We will discuss interest payments below.

What are the conditions that would enable this process to proceed? First, the firm must believe that the purchase of the services from the public will generate sufficient income flows to retire the debt (including interest) and provide the expected rate of profit. The bank will engage in the transactions described above only when the interest it receives on the short-term debt (CP) exceeds the interest it pays on time deposits held by the public, plus the interest it pays on borrowed reserves. The public sells its services if the price offered is sufficient to induce it to hold financial assets (that can be used to purchase consumer goods) rather than real assets (time to do something else). The public is then induced to exchange demand deposits for time deposits, and then for long-term securities by successively higher interest rates.

A MODEL WITH PRODUCTION AND CONSUMPTION

In this section, we look at examples that include production and consumption. We retain the assumptions above, but now assume that a firm purchases labor services from the public that are used to produce goods. For simplicity, we assume that two firms exist, one of which produces only intermediate goods (investment goods) used in the production process, while the other only produces consumption goods for sale to the public. The first is the investment sector firm, whose output is investment goods (I). The second is the consumption sector firm, and its output is consumption goods (C). For simplicity, it will be assumed that the production of each \$1 of consumption goods requires an input of \$0.90 in labor services and an initial \$0.10 in investment (capital) goods. Investment goods depreciate at a rate of 10 percent per year.

Assume the consumption goods firm will produce \$1000 of consumption goods. This requires an outlay of \$900 for services from the public, and \$100 in intermediate goods from the investment sector. (See Figure 9.4, Step 1.) To simplify matters, assume the investment sector firm uses only \$100 of labor

Step 1: Bank Loan Finances Hiring and Production

Commerical Bank	
Asset	Liability
+CP 100	+DD 100

Investment Sector	
Asset	Liability
+LS 100	+CP 100

Public	
Asset	Liability
+DD 100	+100 NW

Step 2: Investment Sector Sells Investment Good to Consumption Sector

Commerical Bank	
Asset	Liability
+PB 100	
-CP 100	
PB 100	DD 100

Investment Sector	
Asset	Liability
-I 100	-CP 100

Consumption Sector	
Asset	Liability
+I 100	+PB 100
I 100	PB 100

Step 3: Bank Lends to Consumption Sector to Finance Production

Consumption Sector	
Asset	Liability
+LS 900	+CP 900
I 100	PB 100
LS 900	CP 900

Public	
Asset	Liability
+DD 900	+NW 900
DD 1000	1000 NW

Commercial Bank	
Asset	Liability
+CP 900	+DD 900
PB 100	DD 1000
CP 900	

Step 4: Consumption Sector Produces Good and Sells to Public

Commerical Bank	
Asset	Liability
-CP 900	-DD 910
-PB 10	
PB 90	DD 90

Consumption Sector	
Asset	Liability
-I 10	
-LS 900	
+C 1000	+NW 90
-C 1000	-CP 900
+DD 1000	-PB 10
-DD 910	
DD 90	PB 90
I 90	NW 90

Public	
Asset	Liability
+C 1000	
-DD 1000	
C 1000	NW 1000

Figure 9.4 Bank makes loan to finance investment

services to produce investment goods. The commercial bank purchases commercial paper from the investment sector and issues the deposit that ends up in the account of the public.

In Step 2, the investment sector firm produces the investment good and sells it to the consumption sector that issues a long-term private bond. Note that the demand deposit received by the investment sector is debited as it retires the commercial paper, but the public retains a \$100 deposit from the sale of labor services. (Note: balance sheet changes are shown above the dotted line, while final positions are shown below the line.) In Step 3, the consumption sector

firm issues \$900 of commercial paper to the bank to purchase \$900 in labor services from the public to produce consumption goods (C), financed through the issue of short-term private debt to banks. (Note: we will ignore profits in the investment sector.)

The final position of Step 3 is shown below the dotted line and includes the results from the previous transactions. The public now holds \$1,000 of demand deposits (Net Worth has risen by selling labor services, \$100 to the investment sector and then \$900 to the consumption sector). The commercial bank holds commercial paper and bonds worth \$1,000.

In Step 4, the consumption sector produces the consumption goods, using the \$900 in services purchased from the public and the \$100 investment good. Only 10 percent of the investment good is depreciated in the production process, so \$90 worth of the investment good will be left at the end of the period. The public spends its entire \$1000 demand deposit (accumulated wages) on consumption. (The sale of the consumption good is shown below the first dotted line; the final result is shown below the second.) The consumption goods sector retires all its short-term debt, and 10 percent of its long-term debt (the term of the long-term debt is assumed to equal the economic life of the investment good).

The net result can be summarized as follows: the public has provided \$1000 in services: \$100 in labor to the investment sector and \$900 in labor to the consumption sector. It has used its income to purchase \$1000 in consumption goods. The investment sector produced an investment good using labor services of \$100, financed through the sale of short-term debt to the banking system. This short-term debt is retired when the consumption sector firm purchases the investment good. The purchase is made by issuing long-term bonds to the banking system, assumed to have a term of 10 years. The consumption sector uses the investment good and \$900 in labor services to produce consumption goods that are sold to the public. Upon the sale of the consumption goods, the consumption sector can retire all of its short-term debt (issued to cover the cost of labor services) and meet its annual commitment on long-term debt, \$10. The consumption sector is left with a demand deposit of \$90 and with \$90 of the investment good (depreciation was \$10).

At this point, the consumption goods sector could retire the entire remaining long-term debt. Alternatively, part of the demand deposit (\$90) could be used to provide interest payments to banks and/or profits to the investment sector (which have been ignored). Note that the net worth of the firm has increased by \$90, which is equal to the net investment (gross investment less depreciation). In this case, net profits also equal net investment: \$90. This is the well-known Kalecki result: net profits equal net investment. Why? Because wages spent by the workers in the investment sector are spent on consumption, while the cost

of producing consumer goods is equal to the wage bill in the consumption sector plus depreciation of the investment good.

If the consumption goods firm decides to liquidate its debt (using the demand deposit to retire the long-term bond), then its net profit exists only in illiquid form (the capital good). We have assumed the consumption sector retires its short-term debt of \$900, but there is no need for it to retire this debt, and it is unlikely that it would. The consumption sector firm has issued short-term debt to hire labor so that the capital it has purchased can be operated over a number of years to generate an income flow. Thus, in each subsequent period, it is likely that the short-term debt will be rolled over as a revolving wage fund. As long as the firm remains creditworthy and makes interest payments on the short-term debt, the bank will allow it to renew this “revolving fund of finance” each period. Similarly, the investment sector firm may utilize a revolving fund of finance since it, too, is a going concern.

Thus, we would expect to find, over the time period required for an entire production and consumption cycle, a close relationship among the total short-term debts issued by firms (commercial paper in this case), the demand deposits issued by the bank that enable the firms to pay wages to the public, and consumption goods purchases. Since the velocity of demand deposits paid as wages is assumed to be one, since the marginal propensity to consume out of worker income is one, since workers are paid once each period and they spend their wages once each period, and since capitalists (firms) do not consume, it must be true that: $\text{Wages} = \text{commercial paper} = \text{demand deposits} = \text{consumption}$.

Of course, a change in any of the assumptions would change these relationships.

We can make several points. First, the production of an investment good has resulted in profit income, just as it has resulted in the production of goods that are not consumed directly. Investment goods are consumed indirectly and only gradually as they depreciate. Production of investment goods allows firms to accumulate these non-consumable goods to be used in further production. Second, the production of investment goods produces income (accumulated in demand deposits) that is available as gross profits (or saving—as we assumed capitalists do not consume), which can be used to retire the debt or to engage in further production. Third, payment of interest to the banks would come out of profit income—that is, income generated by the production of investment goods, which in turn creates wage income that is spent on consumption goods. Interest is a deduction from gross profits. Finally, if banks paid interest on deposits, that would redistribute their interest income (generated by investment) to depositors. Again, such interest income would be a deduction from profits. All of these results are similar to those reached by the Circuit approach (see Chapter 6).

Although we have not explicitly incorporated real chronological time in these models (the time period was long enough to allow the full cycle to be completed), the results would not be significantly different if time were brought into the analysis. In a model with time, wages as a stream of income would not be equivalent to the accumulation of demand deposits, profit income would not be equivalent to demand deposits accumulated by firms, and so on. In other words, a smaller amount of *money stock* (demand deposits) can support greater flows of wage and profit income. Adding a time dimension allows velocity to diverge from one. Once time has been admitted, balance sheets differ from income accounts in these examples.

Clearly, credit expansion that allows the production of investment goods to proceed generates the funds necessary to allow the retirement of the debt. Prior accumulation of funds (savings) that are then intermediated through banks is unnecessary for the investment process. Every expenditure of a demand deposit that has been created through the expansion of a bank's balance sheet results in a receipt. Neither is exogenous money necessary, since endogenously created credit is sufficient to allow production to increase.

In an expansion, a simple expansion of private balance sheets suffices. Below we discuss what happens when liquidity preference rises. Briefly, because the expansion is based on credit, if economic units decide they do not want to hold privately issued credit, the system will collapse. For example, returning to the last balance sheet, assume that the bank decides to liquidate the long-term bond to increase cash reserves. The only cash that exists in the system is the demand deposit liability of the bank.

There are two possible constraints to growth in this model: finance and labor services.⁶ We have assumed an infinite supply of labor services, so this is not an operative constraint. The other constraint is finance that is limited only by the willingness of economic agents to accept private liabilities. Neither saving nor exogenously supplied money is necessary.

Certainly, exogenous money (for example, dropped from helicopters) would also enable the system to expand. However, saving would not. As Terzi (1986–87) has shown, prior saving cannot be the basis of an expansion in an integrated market economy. This is due to the obvious fact that in an economy where production is for the market, income and expenditure must be equal. Any attempt to save income must mean a decline in expenditures, and therefore, in income. If spending is income-constrained, then it cannot grow. Only in a non-market economy can saving be a source of further expansion—for example, a yeoman farmer can save a portion of the corn harvest for next year's planting. In an economy in which a large portion of production is not for the market, thus, where monetary production is not the basis of the economy, saving can be the basis for investment. This "primitive accumulation" can occur because balance sheets are not highly interconnected and the saving leads directly to an

investment in corn production. Thus, in a non-market economy, reduction of consumption can allow expansion.

However, as production increasingly occurs not to directly satisfy consumption needs, but to be sold on the market, saving becomes less important as a source of investment. Once a monetary economy develops and saving takes the form of monetary savings, saving cannot be a net source of funds for investment—Keynes's paradox of thrift rules.⁷

EXTENSIONS OF THE BALANCE SHEET APPROACH: RESERVES AND LEAKAGES

Several simplifications were made above to facilitate analysis. For example, reserve requirements were ignored in our example of financing production. The quantity of reserves required will depend on the length of time needed for each of the transactions involving demand deposits. Banks must obtain reserves or engage in reserve-economizing liability management to meet legal reserve requirements. As Moore argues, the central bank must accommodate the desire for reserves—we can think of them as horizontally supplied. While banks can economize on reserves, they cannot create them. Legally required reserves do not constrain banks any more than banks are constrained without legal requirements—the central bank can determine the price but cannot leave them short. If the central bank pays interest on reserves, it can leave excess reserves in the system, but a shortage would cause the market rate to rise above the target.

Second, we have ignored a) leakages of demand deposits into currency, and b) saving out of wages. Leakage into currency reduces bank reserves, so it increases the quantity of reserves banks must obtain. For example, if the consumption sector firm in the final step above decides to exchange its \$90 of demand deposits for currency, the banking system will need \$90 of reserves. However, the effects on the system are greater if the public chooses to hoard some of its income. If the public decides to save \$90 of its income, then consumption expenditures are only \$910. Now the consumption sector firm does not receive any profits as it can just cover required payments on short-term debt (\$900) and long-term debt (\$10). If the public holds its savings in deposits, then the banking system is not affected: the \$90 in demand deposits are held by the public rather than by consumption sector firms.

Thus, these models confirm Keynes's claim that it is irrelevant to the banking system whether the public consumes or saves—that is, he said that consumption is as much a source of loanable funds as is saving. However, if the public holds currency, then the banking system is deficient in reserves. Thus, the balance sheet results are consistent with Keynes's claim that only hoarding, and not consumption, can reduce the quantity of funds available to banks.

If desired saving exceeds \$90, then the consumption sector firm is unable to cover production costs (wages plus payment on long-term debt). This would force the firm to roll over debt commitments, which might affect production plans. Note also that if the public saves, then consumption and profits are reduced by an equal amount. This is a well-known result from Kalecki's profit equation: saving out of wages reduces profits.

Third, we have assumed that the term of long-term debt equals the life of the investment good. However, aggregate profits (accumulated as demand deposits) are sufficient to retire the entire long-term debt at the end of the first period, so the term of the private bond is irrelevant. The point is that sufficient income is generated to retire all debt (although not to pay interest). If the model were expanded to include several firms, income may not accrue to the firm that has issued the debt. While aggregate profits are sufficient to retire the debt, the firm that has purchased the investment good will likely only capture a portion of the aggregate profits, so it is not able to retire all the debt in the first period. This firm expects to use the investment good in production over a number of years, receiving a portion of the aggregate profits each year so that the debt can be gradually retired. In the aggregate, profits do equal the net credit outstanding, so that all debts *could be* retired, but whether they will be retired depends on the distribution of profits among firms. But retiring all debts is neither necessary nor likely since firms will maintain a *revolving fund of finance*, so that short-term debts will be rolled over to maintain a wage fund. Neither is it desired by banks—they are happy to hold earning assets.

We have arbitrarily assumed that the investment good depreciates at a rate of 10 percent per period. If the investment good fully depreciates in one period, then capital is not accumulated, and no net wealth is created for firms. However, even in this case, aggregate income is sufficient to retire aggregate debt. Since it would be unrealistic to assume that one consumption goods firm could recapture the entire gross profit of one period, debt retirement at an individual level would be quite difficult. But if the economy were expanded to include a great number of firms, each of which is carrying on production by purchasing investment goods and labor services, it is easy to imagine that each individual firm could (but might not) capture sufficient gross profits to retire debts.

EXPECTATIONS, UNCERTAINTY, AND INVESTMENT FINANCE

The simple examples presented above give some insight into the investment finance process. Investment funded through credit expansion generates sufficient income at the aggregate level to retire debt. Credit expansion also

generates sufficient deposits at the aggregate level to enable banks to capture funds to replace those being withdrawn as loans advanced are utilized by borrowers.

However, these models are too simple. Even though aggregate income and aggregate deposits are sufficient, there is no guarantee that individual firms and banks can capture the funds needed. At the aggregate level, this requires that all income is spent so that it returns to the circular flow. Individual firms must obtain a share of the aggregate income to cover costs, repay debt, and obtain a normal profit share. Individual banks must capture deposits to replace reserves as demand deposits are drawn down. This requires that all created funds remain within the banking system (so that no demand deposits are converted into currency or go abroad).

An investment project is undertaken with the expectation that future income flows will enable the future stream of payment commitments to be met. However, aggregate profits (and saving) depend on aggregate investment. For example, assume that in the second period, the consumption sector firm uses the existing capital assets to again produce consumption goods, again purchasing \$900 labor services. However, assume there is no new investment—in fact, investment is negative this period due to depreciation. In Figure 9.5 we examine the impact on profits.

In Step 1, we have assumed the consumption sector used its \$90 demand deposit that remained from the previous period and issued short-term debt of \$810 to purchase \$900 of labor services. Now the consumption sector produces consumption goods, which causes the capital good that was purchased last period to depreciate another \$10. In Step 2, the consumption sector sells the consumption goods produced to the public for \$900.

The consumption sector sells \$900 in consumption goods to the public, receiving a deposit of \$900. It then retires the short-term debt (\$810) and makes a payment on its long-term debt (\$10), leaving it with a demand deposit of \$80 and a capital good worth \$80. Again, the demand deposit is sufficient to retire all the long-term debt. Note, however, that no net profits have been accumulated: the consumption sector holds a demand deposit of \$80 at the end of this period, while it began the period with a deposit of \$90 (carried over from the previous period). Net wealth has declined by \$10, and net investment is -\$10 due to depreciation, which is equivalent to aggregate profits. Thus, aggregate profits depend on investment. Note that this also represents a reduction of saving by \$10 as the consumption sector firm's demand deposit has fallen from \$90 to \$80.

As Minsky (1986) argues, investment this period is undertaken only if investment is expected to occur in the next period. Otherwise, aggregate profits in the next period would not be positive. In Kalecki's simple model, aggregate profits equal investment plus capitalist consumption (out of profits). In

Step 1: Bank Makes Loan to Finance Production of Consumption Good Sold to Public

Commerical Bank		Consumption Sector		Public	
Asset	Liability	Asset	Liability	Asset	Liability
+CP 810	+DD 810	+LS 900	+CP 810	+DD 900	+NW 100
		-DD 90			
CP 810	DD 900	GS 900	CP 810	DD 900	NW 900
PB 90		I 90	PB 90		
			NW 90		

Step 2: The Consumption Sector Sells the Good to the Public and Retires Debt

Commerical Bank		Consumption Sector		Public	
Asset	Liability	Asset	Liability	Asset	Liability
		-LS 900			
		-I 10			
		+C 900	-NW 10		
		-C 900		+C 900	
		+DD 900		-DD 900	
-CP 810	-DD 820	-DD 820	-CP 810		
-PB 10			-PB 10		
PB 80	DD 80	DD 80	PB 80	C 900	NW 900
		I 80	NW 80		

Figure 9.5 What happens to profits if there is no net investment

this example, since net investment was negative, profits can be positive only if capitalists consumed—adding to the consumption spending by the public by drawing down the \$80 demand deposit (or by going into debt), capitalists could sell part of the consumption goods to themselves, and thereby realize profits because total spending on consumer goods would be \$980, allowing a profit of \$80.

In Keynesian theory, investment is primarily a function of profit expectations. As the future is unknowable, these expectations are subject to wide fluctuations. An individual firm cannot know whether its investment project will generate the stream of returns necessary to meet the payment commitments undertaken to finance the project. The individual firm must be concerned with the demand for the product it plans to produce. This demand is at least partially a function of the level of aggregate demand, which determines aggregate income. Thus, expected aggregate demand may also enter the firm's investment decision (big firms have research departments that project aggregate demand as well as demand for the firm's output). Aggregate demand, in turn, is a function of investment and other autonomous expenditures. Investment proceeds when firms are sufficiently confident about the future. Investment, in

turn, generates the aggregate income necessary to enable investment projects to succeed. Expected success, not success itself, is the necessary condition for investment to occur. Since credit can finance investment, it is not necessary for firms to accumulate funds (for example, profits on past successful projects) before proceeding with a project. However, failure to achieve profits that were expected will negatively affect expectations, investment, and thus aggregate profits in the next period.

Similarly, banks will individually create credit only if they expect that they will either retain deposits or attract new deposits and make a profit on the interest rate differential between the loan rate and the deposit rate. Although sufficient deposits are created in the aggregate to cover every extension of credit, individual banks must ensure that their credit extension is covered by reflux or that they are able to issue liabilities in wholesale markets. Leakages of deposits into other banks can be returned via interbank loans, while leakages of deposits into currency (or, in an open economy, out of the country) must be made up through central bank injections of reserves (discount window lending, or open market purchases) or by borrowing from foreign banks.

Banks do not need to intermediate “loanable funds” to make loans. If banks share the profit expectations of prospective borrowers, they can create credit to allow the project to proceed. If the project proves successful, the credit can be retired—although banks generally want outstanding credit to grow, not for credit to be retired. Instead, they want lending to be profitable, and they also need to maintain a sufficiently liquid position to satisfy the liquidity preference of the public. If the public prefers to hold highly liquid assets (for example, demand deposits), banks may have to pay a higher interest rate to attract deposits into less liquid forms (such as time deposits that could be less volatile and do not carry a reserve requirement). Banks must ensure that the interest rate they charge on the loans they make will cover the cost of retaining deposits over the term of the loans, with a margin to cover default risk and to provide a normal return on bank equity.

As discussed in Chapter 7, an increase in liquidity preference can be related to a decline in profit expectations: firms are not willing to invest, which leads to a decline in aggregate demand and profits, making it difficult to meet payment commitments. An increase in liquidity preference can also be related to the public’s preferences for liquid deposits and currency, so banks must obtain reserves to meet reserve requirements on demand deposits or to meet a drain on deposits (resulting from a conversion of demand deposits into currency). If the Fed does not intervene (by providing reserves), banks may try to liquidate assets to obtain liquid reserves. However, this is impossible in the aggregate and—at the extreme—may generate a general run to liquidity.

A BALANCE SHEET APPROACH TO GOVERNMENT SPENDING

Above in Chapter 8, we addressed the MMT approach to government spending. In this section, we will go through the balance sheets. As we saw, in modern economies, governments always spend and tax through their central banks working with private banks. Hence, there are central banks and private banks standing between the government and the public, which complicates the operations to some extent.

We will begin with a government payment to the public, purchasing labor services (although transfer payments—such as social security pensions—operate the same way). We will follow the procedures adopted in the US, although other developed countries follow similar ones. The treasury has a deposit account at the central bank (Treasury Government Account, TGA) and deposit accounts in private banks (Tax and Loan Account, TLA).⁸ Treasury deposits both tax revenue as well as revenue from bond sales in the TLA; however, it moves deposits from the TLA to the TGA as it spends. Central banks generally purchase treasury bonds only in secondary markets (in our example, from private banks), so new issues in the primary markets are purchased by banks (in the USA, by special dealer banks).

Figure 9.6 shows the relevant balance sheets (we will let the public sector be comprised of both households and firms). We adopt abbreviations from above and add Tax to indicate taxes owed to the treasury, Cur stands for treasury currency, and Notes stands for central bank notes. We will assume that the central bank uses open market operations—buying and selling treasury bonds—to supply and drain reserves, ignoring lending at the discount window (which, in practice, accounts for a very small quantity of reserves).

Note that taxes owed are counted as a treasury asset and a public liability. The treasury's net worth is difficult to calculate and potentially limited only by the maximum tax liability it can impose. Obviously, raising the tax liability increases the treasury's net worth and reduces the public's net worth. We ignore taxes on banks; in principle, according to the theory of public finance, all taxes are paid by people—whether owners of firms or employees.

Treasury		Central Bank		Private Bank		Public	
Asset	Liability	Asset	Liability	Asset	Liability	Asset	Liability
TGA	Bg	Bg	TGA	Bg	TLA	Bg	Tax
TLA	Cur		Res	Res	DD	DD	NW
Tax	NW		Notes		NW		
			NW				

Figure 9.6 Balance sheets with a government sector

Let’s assume the treasury wants to purchase labor services from the public in the amount of \$100. It first moves deposits from its TLA at a private bank to its TGA at the central bank. The central bank credits the treasury’s account and debits the bank’s reserves. See Figure 9.7.

In the second step, the government purchases labor services. When the treasury spends, bank reserves are restored, and the private bank credits the deposit account of the public. From the perspective of the private bank, the result is that the public holds a deposit, rather than the treasury. From the perspective of the public, its labor services have been monetized, and its net worth increased; the treasury’s position is the reverse: it has exchanged a monetary asset for labor services.

Let’s see what happens to balance sheets when the treasury receives tax payments (Figure 9.8).

In the first step, the treasury’s TLA is credited by the tax payment, and the private bank debits the taxpayer’s account. Compared with the previous example, in which the treasury paid the public, the taxpayer’s net worth has fallen. However, once a tax is imposed, that is a liability that immediately reduces the taxpayer’s net worth—payment of the tax simply eliminates both the liability and the deposit. The treasury’s position is the opposite: the tax liability has been monetized. The point is that the imposition of a tax liability lowers the net worth of the taxpayers, and payment of the tax simply reduces the liquid position of the public.

In the second step, the central bank credits the TGA and debits the reserves of the private bank. It is important to note that until the treasury places a call on the TLA deposits, bank reserves are unaffected by tax payments. Indeed, this is why the USA created the tax and loan accounts—to prevent tax payments from having undesired impacts on bank reserves. In a banking system in which banks operate with minimal excess reserves, tax payments would place pressure on banks and result in rising pressure on overnight interbank lending rates as banks losing reserves would try to cover the loss by borrowing reserves. Since central banks target overnight interest rates, the central bank would have to provide reserves—either through open market purchases or through discount window lending. The treasury delays moving deposits to

Treasury		Central Bank		Private Bank		Public	
Asset	Liability	Asset	Liability	Asset	Liability	Asset	Liability
-TLA 100			-Res 100	-Res 100	-TLA 100		
+TGA 100			+TGA 100				
-TGA 100			-TGA 100	+Res 100	+DD 100	+DD 100	+NW 100
+LS 100			+Res 100				

Figure 9.7 Impact of treasury spending

Treasury		Central Bank		Private Bank		Public	
Asset	Liability	Asset	Liability	Asset	Liability	Asset	Liability
-Tax 100					-DD 100	-DD 100	-Tax 100
+TLA 100					+TLA 100		
-TLA 100			-Res 100	-Res 100	-TLA 100		
+TGA 100			+TGA 100				

Figure 9.8 Treasury receives tax payments

the central bank until it is ready to spend, which restores bank reserve positions. If the treasury carefully times the movement of deposits to the TGA, it can minimize the reserve effect. And this is precisely how treasuries and central banks operate: coordinating their operations to minimize reserve effects.

The problem is reversed when the treasury spends—putting reserves into the system. If there is a lot of spending over a period in which few taxes are received, banks will wind up with undesired reserves. Trying to lend them in the overnight market will drive the interest rate down and trigger central bank operations to reduce excess reserves. The main way to do that is to engage in open market sales. But if the treasury can time the movement of deposits with its spending, the impact on reserves is minimized.

Note this is precisely the opposite effect of the “crowding out” thesis, which claims that government spending pushes interest rates up. In reality, tax payments put upward pressure on rates, while spending pushes rates down; net deficits place downward pressure on overnight rates. These pressures are relieved by careful coordination and procedures adopted by the treasury and central bank.

As discussed in Chapter 8, treasury spending tends to be spread out during the year, although many tax payments are bunched around quarterly due dates. And if government spending typically exceeds tax revenue over the course of the year—as it does in most countries—the treasury’s net spending will always be adding reserves to the banking system. Over any short period, the central bank can sell off its government bond portfolio, but it could run out. This is why the treasury must sell bonds—to prevent excess reserves from accumulating. To minimize impacts, the quantity of new bonds issued should be close to the budget deficit over the course of the year. Empirically, this is what we find and would expect to find even if the central bank were not prevented from so-called “monetizing” the debt by the prohibition against buying government bonds in the primary market.

In Chapter 8, we discussed alternatives to bond sales to deal with a situation in which both the TGA and TLA are short of funds. One possibility is to allow the treasury to overdraw its account—effectively running a negative balance at the central bank. This would seem to be reasonable—households and firms routinely receive overdraft protection that allows checks to clear even

if they have insufficient funds. The treasury is certainly more creditworthy than households and firms—it is a government agency, and as we have emphasized throughout this book, a sovereign government cannot run out of money. However, central banks are often prohibited from doing so—or, perhaps, they might refuse to do so. What can the treasury do in that case?

There are two options available given the institutional setup we are assuming here, which is based on the USA case. The first option is to sell bonds directly to the central bank, and the second is to sell them to dealer banks. The first option is not normally allowed in the case of the USA; however, many nations do allow it. We will analyze this option first as it is straightforward. Then we turn to the second case, which is normal USA practice. Note that primary bond sales in the USA are handled by dealer banks, who then sell them into financial markets where commercial banks can buy them. Nothing is lost, however, by showing sales directly to commercial banks—we will show what happens when the bonds are bought by banks that offer TLA deposits, which is a much bigger set of banks.

Figure 9.9 shows that the TGA is credited when the Treasury sells bonds to the central bank. The treasury then draws down the TGA and its spending leads to credits to private bank reserves and to the deposit accounts of recipients. Next to allowing overdrafts, this is the simplest way to allow the treasury to spend.

Note that the private bank now has \$100 of reserves—if it had the desired quantity of reserves before the treasury spending, it now has excess reserves. They will be offered in the overnight market and put downward pressure on the interest rate. To reduce that pressure, the central bank will sell bonds. The bonds purchased by the central bank in the first step above will be sold to private banks—this debits bank reserves, offset by the purchased government bond (the central bank’s balance sheet would show a debit of its bond holdings and of its reserve liabilities). Only if the central bank wants to keep excess reserves in the banking system will the bonds remain on the central bank’s balance sheet.

Let’s turn to the second option: selling bonds to banks that offer TLA accounts. Banks normally lose reserves when they buy government bonds.

Treasury		Central Bank		Public	
Asset	Liability	Asset	Liability	Asset	Liability
+TGA 100	+Bg 100	+Bg 100	+TGA 100		
-TGA 100			-TGA 100	+DD 100	+NW 100
+LS 100			+Res 100		

Figure 9.9 Treasury sells bonds to central bank

However, there is a special arrangement for banks that offer deposit accounts to the treasury. As discussed in Chapter 8, the “L” in “tax and loan accounts” stands for “loan,” which takes the form of purchasing a treasury bond—as if the bank is “lending” to the treasury. This is not really a loan, and as we saw, the treasury cannot really spend the deposit. This practice was developed to relieve pressure on banks while allowing the treasury to ensure it can always get credits to its deposits at the central bank. So the bank simply credits the TLA deposit when it buys a bond—it does not lose reserves with the purchase. See Step 1 of Figure 9.10.

When the treasury is ready to spend, it puts a “call” on its TLA accounts, letting banks know what is coming. By timing the call with the treasury’s spending, the reserves are simultaneously removed and restored to the banking system. This might sound like some kind of gimmick to get around the constraint on overdrafts of the TGA—and indeed, that is the effect—but really, it was developed to avoid unwanted impacts on bank balance sheets. The problem is that if one does not understand the operating procedures, it seems that the bond sales are, indeed, a borrowing operation necessary to finance government spending in excess of tax revenue.

To conclude this discussion, we can see what happens in Step 2 when the treasury moves the deposits to the TGA accounts at the central bank, leading to a debit by the central bank of the private bank’s reserves, in the equal amount of \$100. The bank is temporarily short of reserves, but they are restored to the banking system when the treasury spends, as shown in Step 3. The treasury purchased goods or services from the public, which now has a demand deposit

Step 1: Treasury Sells Bonds

Treasury		Private Bank	
Asset	Liability	Asset	Liability
+TLA 100	+Bg 100	+Bg 100	+TLA 100

Step 2: Treasury Moves Deposits to the Central Bank

Treasury		Central Bank		Private Bank	
Asset	Liability	Asset	Liability	Asset	Liability
-TLA 100			-Res 100	-Res 100	-TLA 100
+TGA 100			+TGA 100		

Step 3: The Treasury Spends

Treasury		Central Bank		Private Bank	
Asset	Liability	Asset	Liability	Asset	Liability
-TGA 100			+Res 100	+Res 100	+DD 100
+LS 100			-TGA 100		

Figure 9.10 Treasury sells bonds to banks

at the private bank. The bank losing reserves is not likely to be the one gaining reserves—but excess reserves are redistributed through the overnight inter-bank lending market (fed funds in the USA).

Before central banks were created, governments spent their own currency directly into the economy and then redeemed their currency in tax payments. The procedure was simple and easy to understand. It was obvious that governments had to spend first, then tax. Governments could also accept private liabilities, such as banknotes, in payment of taxes. In some cases, they issued liabilities to banks to obtain the notes. The first central banks were created to finance government spending—treasuries would issue debt held by the central banks, which would provide their notes.

As discussed in Chapter 3, private banks began to “pyramid” on central banks, using central bank notes for clearing. Eventually, treasuries largely stopped issuing their own notes and relied on the central bank to issue notes. Private banknotes were gradually replaced with bank deposits and central bank reserves. Treasuries began using checks for most payments, written against accounts held at the central bank, which would clear accounts at private banks. Treasury bonds had always assumed a bigger role in the economy in times of war, but as governments grew relatively larger and took on greater responsibilities, bond sales became increasingly routine—as did larger deficits, with annual spending usually greater than taxes. The impacts of fiscal operations on the banking system became more important. Central banks discovered the reserve effects of their purchases and sales of bonds in open market operations. Operating procedures evolved over the entire 20th century, largely focused on minimizing the impacts of fiscal operations on banking reserves. This was necessary for otherwise the central banks would have trouble stabilizing the overnight interest rate.

Today, of course, most payments are “electronic”—debits and credits to accounts kept at private financial institutions and the central bank. Unfortunately, the increasingly complex procedures obscure the reality of government spending—which at its base has not really changed. Government spending puts “currency” in the form of central bank reserves into the economy; tax payments take “currency” in the form of bank reserves out of the economy—in redemption. Logically, the reserves must be put into the economy before they are taken out. Spending must create the income that can be taxed away. That logic has not changed.

We have shown that bond sales (whether by the treasury or by the central bank) function to drain excess reserves. This view builds upon Lerner’s second law of functional finance: “the government should borrow money only if it is desirable that the public should have less money and more government bonds.” (Lerner, 1943, p. 40) While Lerner’s use of the term “borrow money” is misleading, he was correct to argue this has more to do with the public’s

portfolio preferences. More specifically, bond sales are designed to substitute a higher interest-earning government liability for non- or low-interest-earning bank reserves and is properly thought of as a monetary policy operation rather than a fiscal policy operation.

If government spending can be financed through the creation of sovereign currency, then why does the government sell bonds? Of course, governments often believe that they must sell bonds to borrow the funds necessary to finance spending in excess of tax revenue—just as firms purchasing investment goods must sell bonds to fund their position in the assets. However, this is an illusion, as the spending logically comes first. The investment goods are produced using short-term finance, while the ultimate purchaser sells bonds to fund ownership of the capital good. In the case of government spending, whether it takes the form of purchasing goods or services, or making transfer payments, it is always financed by a credit from the central bank to a private bank's reserves; the private bank credits the deposit account of the recipient. This is true whether the government runs a surplus, a balanced budget, or a deficit over the period (quarter or year).

In the next section, we will examine the case for eliminating bonds.

BONDS SALES AND INTEREST RATE POLICY

Should the government continue to issue government bonds? The Fed now pays interest on reserves, so it can just leave excess reserves in the system—they are equivalent to government bills, typically 30-day short-term debt, whose interest rate closely tracks the overnight rate target. Politicians, many economists, and the public at large find government debt to be very scary. They fear the debt is too big; it forces the government to pay a lot of interest; and maybe the government will go bankrupt? Perhaps it would be best just to dispense with selling debt? Let's examine the pros and cons.

We can identify three legitimate reasons for selling sovereign government bonds given current law and institutional arrangements: 1) interest rate management; 2) legal constraints; and 3) to satisfy the public's portfolio preferences. Let's begin with the first, which we already touched on earlier.

As discussed above, all government spending is initially financed through the issuance of sovereign currency—today, in the form of central bank reserves credited to private banks in an amount equal to government spending. Tax payments reverse that—reserves are debited. If the budget is balanced, there is no net increase in banking system reserves. For reasons discussed earlier, however, sovereign governments usually run deficits—meaning that net reserves are continually added and need to be withdrawn by bond sales unless the central bank is willing to let the interest rate fall to zero or to the support rate it

pays on reserves. A functional purpose of bond sales is to provide a higher interest-earning asset to absorb excess reserves.

Let's explore the implications of changing procedures.

1. Treasury Stops Issuing Bonds and Central Bank Pays Interest

What would be the alternative to issuing bonds? The treasury can stop issuing bonds, with the central bank allowing overdrafts. In the USA, the treasury has occasionally issued a *special certificate of indebtedness* to the Fed as it ran overdrafts. If policy were changed to allow this as normal procedure, the central bank would offset reserves it creates with the treasury's certificate representing the treasury's debt to the central bank. The treasury remains in debt, but it is simply internal accounting and can be carried on the central bank's books forever. (Note that inter-government debt is not unusual—the Social Security Trust Fund holds nonmarketable debt of the US Treasury issued over the past four decades during which payroll tax receipts exceeded Social Security payments.) If the goal is to eliminate bonds so that the treasury won't have to pay interest (that would add to a deficit), the central bank will not charge interest or fees on the overdrafts. Bank reserves would increase over time in step with the growth of those certificates.

So long as the central bank pays interest on reserves, the overnight market rate will remain above that floor. Usually, the Fed sets a discount rate above that to act as a ceiling and the fed funds rate remains in the corridor (between the rate paid on reserves and the discount rate). Since the central bank would be paying interest on reserves but not receiving interest on the treasury's overdraft, the central bank would incur accounting losses in perpetuity. While that would not be unprecedented, it would be unusual. But it could go on as long as desired. The central bank can always pay interest on reserves by crediting reserve accounts of banks—whether it is making losses or not.

What are the implications of eliminating bonds in this case? If we take the USA in 2024 as an example, the outstanding debt was approximately \$30 trillion; if instead of bonds, those were excess reserves on which the Fed paid interest near its then-current target rate of about 5 percent, Fed losses on those reserves would run about \$1.5 trillion annually. (Note that the Fed's bond holdings peaked at about \$9 trillion in 2022, after which the Fed began to reduce its portfolio. Much of that debt was low-earning, and Fed losses were substantial because much of the bond portfolio was earning less than the market rate.) Of course, if the Treasury were to pay interest on overdrafts, that could eliminate the Fed's losses, but from the Treasury's perspective, it would be equivalent to issuing short-term bills. The Treasury already has that option—although it typically issues a range of maturities, so pays a higher rate on average than the bill rate.

Alternatively, the central bank could sell bonds to replace reserves with a higher-earning option. While it might sound strange, many central banks do sell bonds—especially when fiscal policy does not create enough treasury bonds to satisfy market demand. This also allows the central bank to continue to manage liquidity by selling bonds of varying terms. Again, the problem is that the central bank will incur even higher interest costs that probably cannot be covered by fees. Central banks do not usually operate at a loss, although they can continue to make interest payments no matter how big the losses are. The question is whether perpetual central bank losses are more politically acceptable than perpetual treasury debt.

2. Eliminating Bonds With ZIRP

Some MMT proponents advocate a policy of permanent ZIRP—*zero interest rate policy*—with a zero interest rate as the target. Eliminating interest on reserves and maintaining excess reserves in the system will ensure that the market rate will hover around zero. In this case, the central bank would neither earn nor pay interest as a result of its financial operations on behalf of the treasury—so central bank losses can be eliminated (the central bank would have some operating costs, but these might be offset by other revenues, such as charging banks for using central bank clearing).

In this case, banks would hold a large quantity of non-earning assets. A significant portion would be held by foreign banks and central banks—about 8 percent of the Treasury bonds were held outside the USA in 2024. As banks are profit-seeking, it is somewhat difficult to see such a situation as sustainable. Indeed, one of the reasons Congress originally allowed the Fed to pay interest on reserves beginning in 2009 is that QE would put massive quantities of reserves into the banking system—and if the Fed did not pay interest on those, a large proportion of bank assets would have been non-earning. Furthermore, the current Basel II supplemental leverage ratio includes bank reserves as an asset against which capital must be held; if banks held a total of \$30 trillion of reserves, they would need massive equity injections to meet the standard—a highly unlikely possibility that would further reduce the profit rate on equity.

Of course, that \$30 trillion of treasury debt was built up over the entire history of the USA, and instead of eliminating it, we might just halt future issues of government debt. In that case, reserve holdings (and Treasury overdrafts) would gradually accumulate over the years as the government runs deficits. Banks (and the Fed) would gradually accumulate assets that reduce profitability. The bank profitability problem is postponed, but not resolved.

We conclude that bond sales are not needed to allow the treasury to spend and the central bank to manage the overnight interest rate, but eliminating

them does raise questions about bank profits and central bank losses in a nation that runs budget deficits but does not issue interest-paying treasury bonds.

3. Treasury Issues a Trillion-dollar Coin

As discussed, under current rules, the Fed cannot allow Treasury overdrafts (and this is a common prohibition among developed nations). Is there another way around the prohibition? Some MMT proponents have advocated a *trillion-dollar platinum coin*—something that appears to be legal. The Treasury can simply issue very high denomination coins (remember, the Treasury can coin money, although it does not currently routinely issue them to finance spending—but some recent coinages have been undertaken with a view to gaining revenue from “seigniorage”) and deposit them in its TGA at the Fed. The Fed would then continue to make payments for the Treasury. Again, the Fed would hold a non-earning asset but would pay interest on reserves, making losses. With ZIRP, income losses are passed along to banks as they will be holding non-earning reserves as assets. The platinum coin resolves the Treasury debt “problem,” but has the same disadvantages as the overdraft solution.

4. Treasury Notes and Central Bank Digital Currency as Alternatives

There are two more radical alternatives to the current system: the treasury could simply return to spending by issuing notes, or the central bank could make payments directly to recipients on behalf of the treasury. Both would bypass the private banking system. The first doesn’t require much explication as that is the way treasuries used to spend: the government simply prints cash. Since cash is not normally thought of as debt, this would eliminate the national government debt (as usually defined) and interest payments on it. It would, however, be quite inconvenient and expose the payments system to risk. Armored trucks would have to deliver cash, and recipients would need to make provisions for safekeeping. Presumably, they would run to their banks to deposit the cash—passing the risk and costs to them. Banks would want to ship the cash back to the government in exchange for reserves. In that case, all the disadvantages of operating a banking system with \$30 trillion of bank reserves discussed above would apply. The central bank could refuse the deposits—but that would remove the government’s money from the private *giro*—a highly undesirable proposition. Government currency would be issued to finance government spending and redeemed in tax payments—with the rest accumulated in private hoards. If banks and central banks refused to accept it, currency would not be very useful in everyday transactions, either.

However, some proponents advocate moving to central bank digital currencies: everyone would have an account at the central bank.⁹ The monthly social security payment could be paid directly by the central bank (on behalf of the treasury) into the retiree's account at the central bank. Tax payments reverse that: the taxpayer's account at the central bank is debited and the treasury's account is credited. Note that the system will work more smoothly if the treasury is allowed overdrafts—otherwise, the treasury will still need to sell bonds to central bank digital currency holders to get a transfer of electronic funds into its account.

If private banks are kept out of the loop, there will be no direct impacts on bank reserves from treasury spending and taxing—however, it is probable that taxpayers and recipients of government spending will be moving funds back and forth between their private bank accounts and their accounts at the central bank. Thus, the central bank will still need to offset reserve effects of fiscal operations. If the treasury is allowed overdrafts so that it doesn't issue bonds, the central bank can issue its own bonds to drain reserves from banks. While this might sound strange, some central banks do this, and it is really nothing more than offering something like a savings account as an option to a checking account. In this way, banks do not end up with low-earning excess reserves, but on the other hand, the central bank will make losses (as discussed earlier) unless it charges sufficiently high fees on the use of digital currency to cover interest on its bonds. From that perspective, this proposal has the same disadvantages as the substitution of bank reserves for government bonds in private bank portfolios.

5. Economic and Policy Implications of Bond Elimination

An important reason for issuing bonds is to provide the public with a portfolio of safe, earning assets of various maturities. With the treasury issuing a range of maturities, the central bank gains some influence over the term structure of interest rates—meaning that it has an additional policy tool at its disposal. Not only does it directly set the overnight rate (and can try to influence expectations of future rates, thereby affecting long-term rates through the expectations channel discussed in Chapter 7), but can also directly influence long-term rates through the choice of maturities it buys and sells. In principle, the central bank can peg the rate on any maturity it stands ready to buy and sell. How much this would affect long-term rates on other kinds of (nonsovereign) debt is uncertain. However, central banks typically buy, sell, and lend against longer maturity debt to manage liquidity. For example, if financial markets try to unload longer maturity debt, the central bank can buy it to stop its price from falling (and prevent longer-term interest rates from rising). (See also Fullwiler 2020.)

As of the end of 2022, just over three-quarters of US Treasury debt was held by the public; the rest was held within the government (various trust funds, for example) and by the Fed. Of the debt held by the public, about two-thirds was held domestically and a third by foreigners—with Japan and China as the top holders (foreign ownership is closely related to the trade balance, with net exporters to the USA and offshore banking centers accounting for most of it). Financial institutions are the biggest holders, with mutual funds, banks, and insurance companies being the biggest among them. Next come state and local governments and pension funds, with savings bonds representing a tiny slice and about 10 percent in the “other” category that includes individuals—who presumably tend to be wealthier than the average American. Bonds are held as a safe asset in private portfolios, and US Treasury bonds are the most important financial asset used as collateral against other debts. Long maturity Treasury bonds help insurance companies and pensions match income flows against expected future payments and help banks and other financial institutions meet capital and liquidity requirements.

If bonds were eliminated, it is possible that foreign holders might try to shift portfolios out of dollar assets since foreign banks and central banks would hold reserves paying low or zero rates (what the Fed would pay on reserves would still be a policy decision of the USA). This could impact the dollar exchange rate—although it could be argued that foreigners would shift from government bonds to private US assets, driving up their price if they still preferred the dollar. Note that if foreigners did buy private US assets, this would put the non-earning reserves back into USA banks.

Domestically, financial institutions and other holders would find alternatives to government bonds—with some buying local and state government debt or private assets. They would take on more risk and have less-liquid portfolios. In the absence of US Treasury bonds, there would be no default-risk-free interest rate greater than the rate paid by the Fed on bank reserves. That might reduce interest income flowing to high-wealth households unless they were willing to take on more risk.

To sum up, ending bond issuance and paying no interest on reserves would have negative effects on bank earnings and on private entities that currently can add safe government bonds to their portfolios. There is also some chance that there could be negative impacts on the exchange rate. Ending bond issuance but continuing to pay interest on reserves would shift indebtedness and losses to the Fed, although, in principle, this is sustainable. The advantage of ending bond issuance is that all the fear of *government debt* might fade away, so long as we ignore the central bank’s debt in the form of reserves.

We conclude that the current system in which bonds are issued might be preferable to one in which excess reserves remain in the banking system when the government runs a deficit. However, the prohibition on treasury overdrafts

as well as on central bank purchases of new issues play no obvious useful role and are a historical relic that dates to the British Parliament's attempt to wrest fiscal power away from the Crown. The USA debt limit is counterproductive as it increasingly leads to political infighting and threatens voluntary default should politicians fail to come to their senses. Rather than trying to eliminate government bonds, it would be far better to explain that both government deficits and debt are normal and beneficial for most sovereign currency nations, and to drop the USA debt limit.

For these reasons, it would be good to eliminate the prohibition on treasury overdrafts, allow the central bank to purchase bonds in the new issue market, and get rid of the debt limit. However, retaining treasury bond issues seems to be the best option for now—providing a safe interest-earning asset with interest paid by the treasury to help cover the costs to banks of running a national payments system. The national debt also provides a safe asset for savers and safe collateral that underlies the financial system. However, policymakers should carefully examine the consequences of allowing central banks to push up interest rates, raising the risk-free reward paid to holders of government debt. Both the interest rate paid and who is allowed to hold government debt should be subjected to scrutiny to ensure the public interest is served.

CONCLUSIONS

In this chapter, we have used balance sheets to examine how private and government spending are financed through credits and debits. What we have emphasized is that the conventional view of banking as *intermediation* is wrong: banks do not first take in deposits that they can lend. Rather, the process of lending creates the deposits—and the spending simply moves the credits and debits among balance sheets. The same is true for government spending and taxing, which lead to a complex series of credits and debits affecting the balance sheets of recipients and taxpayers, banks, the central bank, and the treasury. While complex, we can draw important conclusions from the logic of the balance sheet operations. In a sense, balance sheets provide an important check on our theory of how money works, because the balances have to balance. While everyone has heard of double-entry bookkeeping, when we analyze monetary operations we need at least quadruple-entry accounting because at least two parties are involved and each has an asset and a liability. When it comes to government spending and taxing, we have many more balance sheets involved. Let's summarize the most significant points we have made.

We have seen that short-term credit can finance the wage bill, which flows back to consumption sector firms and replenishes the wages fund or allows the retirement of short-term bank debt. Wages paid in the investment sector generate profits received in the consumption sector (which can be divided among

consumption sector firms, investment sector firms, and banks). Accumulation of investment goods is matched by an equivalent accumulation of financial assets issued to fund positions in them, which take the form of long-term bonds and equity held by the public. Alternatively, banks might accumulate the long-term bonds and allow firms (or the public) to hold demand deposits. This confirms Kahn's (1954) observation that the "money supply" is partially determined by the quantity of bonds the banking system is required to hold. At any point in time, the quantity of demand deposits outstanding is determined by the banking system's holdings of short-term and long-term debt that has not yet been retired.

Exogenous money dropped from helicopters will also allow economic growth. However, in any modern monetary economy, money is privately created rather than exogenously dropped. (It is hard to think of cases of exogenous money drops, although one could argue that COVID relief in the USA came close—with about \$5 trillion in government checks mailed out to most households.) Helicopter money is a gift—what is called a transfer payment, a credit with no liability imposed. In any case, it is clear that while endogenous and exogenous money allow the system to expand, saving cannot do so. It is the willingness to increase spending that allows growth and generates saving. While the quantity of accumulated savings is always sufficient in the aggregate to retire all debts, saving flows are only made possible by accumulation elsewhere of debts.¹⁰

In a credit economy, a constant level of spending requires a constant level of finance in the form of a revolving fund to cover the wage bill. The revolving fund will have to be continually renewed by banks to allow production to continue. An increase in spending requires an increase in the quantity of finance advanced by banks and will then lead to a larger revolving fund that can be continually renewed.

Turning to government spending, it leads to a credit to a bank's reserves and to the demand deposit of the recipient of the spending. Taxes simply reverse that. Logically, taxes are a leakage, like saving. When government spending exceeds tax receipts, the non-government sector experiences a net increase in income. The addition to savings equals the government's deficit. The logic is equivalent to that of private investment: government spending also generates the income that can leak out of the economy—into taxes, domestic saving, and—in an open economy—to a current account deficit (rest of the world saving).

The net savings created by government deficits is initially accumulated in the form of currency—today, bank reserves. The functional purpose of bond sales is to drain reserves, offering a higher-earning and safe, asset. Given current operating procedures in the USA, the deficit (a flow over the fiscal year) will be largely accumulated as a stock of government debt (more or less equal

to the deficit over the period). Payment of interest on the stock of debt adds to the government's spending, hence, increasing the deficit as the debt grows.

While there is no danger that the government would not be able to pay interest on the debt, some worry about default while others argue that it is not necessary for the government to pay interest on perfectly safe debt. We have examined the consequences of eliminating bond sales and of halting interest payments. We concluded that while the current system has disadvantages, it might be the least disruptive option.

NOTES

1. Boulding also recognized the usefulness of a balance sheet approach to money and spending. His early texts (the first edition appeared in 1941) used balance sheets to show how loans make deposits. He also argued that all banks, acting together, could increase the money supply:

A many-bank system is rather like a number of balloons tied together with a string; if one balloon tries to get away from the others the strings will bring it back, but all the balloons can rise together without difficulty. The "string" in the case of a banking system is the loss of cash reserves. One bank expanding loans disproportionately will lose reserves to the other banks, but if all banks expand together they will all lose reserves to each other, which means of course that no bank loses reserves on balance except to the public. (Boulding 1966, p. 107)

2. All assets are valued in dollars or wage units. We assume throughout that there is no inflation. Inflation would complicate the analysis. If wages increase commensurately with general price increases, then it would become easier for firms to repay debts. If wages increase slower than prices generally, then firms may not be able to sell all consumption goods produced. These problems will be ignored by assuming no inflation. Since we are also assuming that the public can supply labor services on demand, labor bottlenecks will not occur. Thus, the assumption of no inflation is not unrealistic for these simple models.
3. It is unrealistic to assume that firms and banks initially begin with zero net worth. The net worth of a firm is an important determinant of its ability to obtain credit. The net worth of a bank is an important factor that makes bank liabilities (such as demand deposits) liquid. Other factors contributing to the liquidity of bank liabilities include portfolio diversification, deposit insurance, and the willingness of the central bank to operate as a lender of last resort.
4. We assume the public does not purchase commercial paper or government bonds.
5. Equilibrium is used merely to indicate a position in which the balance sheet held is the one desired by banks and firms. In the case of the bank, reserves

equal required reserves in equilibrium. Equilibrium in the sense used here has nothing to do with full employment. In fact, we have essentially assumed there is an infinite supply of labor services available.

6. A third constraint occurs if the marginal propensity to consume out of wages declines below unity. This possibility will be discussed in the next section.
7. See Terzi (1986–87) for a proof of this statement using flow-of-funds accounts. See also Wray (1988).
8. Note that the US Treasury has not been using the TLA in recent years as banks have massive excess reserves due to Quantitative Easing, on which the Fed now pays interest so there is no imperative to reduce reserve effects of Treasury operations.
9. Rohan Grey has a more complex proposal for a central bank digital fiat currency that would avoid some of these problems. Private entities (households, firms, and banks) as well as the treasury would have digital accounts at the central bank, so the government could spend by directly crediting the recipient's account. The private and public *giros* would be integrated on the balance sheet of the central bank. Rather than issuing deposits as they make loans, private banks would run overdrafts on their central bank accounts (paying interest) when they make loans (earning interest). The loan recipients would receive credits to their digital accounts (and pay interest on the loans). If the central bank pays interest, this takes the form of a credit to the digital account, or it could issue digital bonds (effectively, just a savings deposit) that pay interest. This would eliminate the need for dealer banks to handle government bonds. It also could eliminate securitization (that was behind the GFC) as banks would hold the loans as assets against their digital overdrafts at the central bank. Finally, the treasury could also run overdrafts on its account, making payments directly with digital fiat currency. There would be no treasury debt to worry about—just overdrafts on its central bank account. The central bank would issue debt in the form of digital credits and would incur interest costs if it agreed to pay interest on accounts. If the treasury pays interest on overdrafts, that will increase its overdrafts and offset central bank interest costs. See Rohan Grey (2019).
10. Foster argues:

Aggregate income cannot be transferred from one time period to another. This means that one generation cannot implement or diminish the income of future generations by expending less or more than its own means. The traditional concern with this issue arises out of the palpable fact that individuals can hoard, and they can incur debts beyond their accumulation of assets. These opportunities to individuals arise out of the fact that an individual can, and often does, buy more than he sells or sell more than he buys. But this opportunity is not available to the entire economy ... sales

and purchases are necessarily equal. ... There is no 'time lag' here. The community at large cannot 'save money'; it can save only by investing, and its savings are constituted by that investment ... [T]he equation between aggregate sales and aggregate purchases forces the instantaneity of the equation between saving and investment. (Foster 1981, pp. 967)

10. Money and the logic of capitalism

INTRODUCTION

As discussed above in Chapters 2 and 3, the development of money predates markets, and the development of markets long predates the development of capitalism. However, the monetized sectors of precapitalist societies were small. In such societies, credit was primarily responsible for financing inventories in the process of transport, consumer expenditures, and spending by the authorities. A wide variety of credit instruments was used to satisfy these purposes. Foreign trade generally was associated with more advanced credit forms, primarily due to the impersonal (and, therefore, risky) nature of such trade. As discussed in Chapter 2, the early money of account was derived from a grain unit such as barley, wheat, or rice, recorded in a variety of ways (tablets, hazelwood tallies, paper) and relatively late in the form of stamped coins. In most cases, both government-issued money and privately issued credit money existed side by side in precapitalist societies.

Once credit money has led to the development of coined state money, the coins sometimes dominate other forms of money in markets as a medium of exchange and means of payment in precapitalist societies. In this chapter, we argue that the situation is far different with the development of capitalism, where the logic of the system requires that private credit money is the dominant form of money and that the use of currency (coin or notes) in exchange or payment is the exception. Using the term from Chapter 6, money must be *endogenous* to the system. Most importantly, capitalism changes the logic of production—from satisfying economic needs to the accumulation of capital in the form of money. The modern financial system evolves to facilitate that purpose—but as we shall see, it also creates problems that the state can help to resolve. The state also uses the monetary system to mobilize resources in pursuit of the public purpose—including to counter problems created by capitalism such as rising inequality, poverty, unemployment, discrimination, and damage to the environment.

Goldsmith argues that it cannot be a coincidence that the modern financial system developed with the industrial revolution. (Goldsmith 1987, p. 4) However, we will argue that the emphasis on the First Industrial Revolution in northern Europe is probably misplaced. The main institutions of capitalism,

and especially its financial system, came out of the New World's plantation economies. For our argument here, where capitalism came from is not so important—rather we want to focus on the unique role played by money in capitalism. Still, we will first review the historical evidence for the origins of capitalism and modern financial institutions.

Before proceeding, let's quickly define what we mean by capitalism. Many definitions would include private ownership of the means of production, production for market, and “free” wage labor—all of which are characteristics of modern capitalism. However, we will emphasize the following features of capitalism:

1. the commodification of labor—labor power (or working hours) can be bought and sold for money—whether free or unfree;
2. a generally growing proportion of social output in terms of the means of livelihood is commodified (sold for money);
3. generation of a social surplus in money form; and
4. accumulation of a portion of the social surplus as capital by the capitalists—who use it for further accumulation.

Or, more succinctly, capitalism is defined by Marx as $M-C-P-C'-M'$ (starting with money capital to purchase the commodity labor power—as well as produced means of production—to produce a commodity to sell for more money) or as a monetary production economy by Keynes, where the goal is to end up with more money than the capitalist started with.

MONEY AND THE RISE OF CAPITALISM

Another look at David Graeber's assessment of the myth that money originated out of market exchange based on barter is a useful place to begin this discussion.

It is so deeply established in common sense, even in places like Madagascar, that most people on earth couldn't imagine any other way that money possibly could have come about. The problem is there's no evidence that it ever happened, and an enormous amount of evidence suggesting that it did not. (Graeber 2011, p. 28)

Graeber's careful analysis demonstrates that barter has never been an important economic activity, that barter exchange did not lead to the development of markets, and that money could not have developed out of barter. Furthermore, our modern capitalist system did not rise out of a pre-existing market system based on “free” trade and labor. While it is true that both money and market exchanges existed long before capitalism, Graeber argues that it wasn't

until the 19th century that we find anything close to what we identify as free labor—workers choosing to enter into contracts to work for wages, and free to quit jobs to seek alternative employment. Instead, there were gradations of indentured labor—with New World-type slavery at the far end of the extreme.

In line with this view, modern scholarship has revived a thesis put forward by Eric Williams in the 1930s that locates the beginning of modern capitalism in New World slavery.¹ Surprisingly, our modern financial systems likewise evolved out of slavery—not out of an earlier European industrial revolution.

It has been common in economics to argue that capitalism emerged out of feudalism in northern Europe, taking the form of small-scale factory production. (See Dobb 1963 for a classic study of the development of capitalism.) Peasants fleeing excessive rent imposed by feudal lords sought jobs in these new factories, earning wages. Certainly, there was some force involved—the *enclosure movement* reduced the ability of peasants to survive by cutting off access to forests or pastureland—but there was also the pull of greater freedom and opportunities in the new cities. (Heilbroner and Thurow 1982) Gradually markets became *freer* as Mercantilism and its protection of exports was abandoned in favor of liberal trade policy advocated by Adam Smith. New World slavery was an inefficient aberration that held back innovation and was finally eliminated in the United States by the Civil War.

Of course, that is a potted version of the economic history that has long been taught, but, importantly, Williams (1944) argued that all of this is on the wrong track. The major features of modern capitalism cannot be found in the small-scale industry of Northern Europe but on the huge slave plantations of the New World. Even Europe's development of factory-based capitalism was built on the profits from the well-known triangular trade: captured African slaves shipped to the New World's plantations to supply cotton and sugar for the factories in Europe that produced consumer goods, some of which were shipped south on the way to pick up more slaves. Modern management techniques as well as careful bookkeeping were developed on the plantations. And modern finance grew out of the slave trade: the bodies of slaves provided the collateral behind the loans that financed the triangular trade to generate profits that financed the growth of capitalism.

Rather than being the laggard, America's South (as well as the plantations in the West Indies and Latin America) was the source of wealth for the nation—and for the development of Europe. The productivity (in terms of profits) of a slave in the New World was orders of magnitude greater than that of the worker in Adam Smith's pin factory. Williams estimated that one person (White or Black) in one of England's colonies produced as much profit as 10 in Britain; every person on a sugar plantation in the New World created 130 times as much profit as a worker at home in Britain. (Even Adam Smith argued that the sugar plantation in the New World was more profitable than any farm

in Europe (Graeber 2006); Williams (1944) argues that the Caribbean was far more important to Britain than the American mainland because of the sugar trade.) It is also important to note that the skills and knowledge of both slaves and indigenous people contributed greatly to success in the New World. The enslavers preferred to kidnap highly skilled Africans who then taught the New World colonists how to grow rice and helped construct the USA Capitol, for example.

The rise of the modern nation-state is similarly tied up with New World slavery. Slavery required supporting institutions to protect private property (largely slaves) and to capture and punish slaves who escaped as well as those who assisted them. The state had to develop modern navies to protect slave ships as well as transport the products of the plantations and factories. Those factories had to innovate to increase capacity to process the New World's almost limitless productivity—thanks to plentiful slaves and natural resources. And, as mentioned, the triangular trade required the creation of modern financial institutions with their instruments such as mortgages and bonds. Wall Street served as a principal center of both commerce and finance based on slavery from its beginning. The wall after which Wall Street was named was built by slaves for the first slave auctions—and in New York City, all Black and Indian slaves had to be bought at that market. An estimated 12.5 million Africans were captured and enslaved, of which 400,000 were sold in America.

The value of slaves in America was greater than the value of all banks, factories, and railroads combined. Slaves were worth far more than the plantations on which they worked.² Their bodies were collateralized with mortgages and bonds that provided the finances to buy and run the plantations, and insurance developed to protect investments in their bodies should they die during transport. Many of the early British bankers came from businesses related to slavery—merchants, traders, and shipowners. The slave trade provided finance for the British munitions industry as well as for Watt and his steam engine. Modern accounting was developed for calculating probabilities and depreciating values. The legal and ideological infrastructure that we associate with an economic system based on private ownership was developed to commodify things and humans, and to establish standards for measuring, grading, and pricing. What we now call “financialization” got its start in creating securities markets, sovereign debt, collateral, and investment opportunities. (Williams 1944, Rockman 2014)

Innovations developed faster ships to reduce losses from deaths of slaves on the long voyages (an estimated 2 million died during the so-called “middle passage” across the Atlantic), and better armaments were mounted on the slave ships to protect the human cargo. The nation that excelled at transporting slaves (Britain as the leader, followed by Portugal and France)³ would inevitably dominate the seas with the best naval vessels and could call on the

well-armed slave ships to join in the fight when necessary. American slavery “made” Britain “Great” even as it developed America’s might. The triangular trade also promoted fishing, the market in dried, salted cod, and the growth of great seaport towns and docks. It fueled rum distilling, metal manufacturers, production of chains, locks, and guns to keep slaves in bondage, and the munitions needed for the mercantile wars and slave revolts. The slave trade was far more important to Britain’s development than were its coal and tin mines or wool, according to Williams (1944).

In the case of the United States—from its beginning as a capitalist nation—slavery was important to the founding of its most important political institutions as well for features of its Constitution. In a twist of irony, it was the wealth created by slavery that enabled and encouraged the Americans to break away from Britain. Forty of the 56 signers of the *Declaration of Independence* were slave owners, many of whom worried that Britain might abolish slavery in the colonies. Ten of the first 12 American presidents were enslavers, and eleven of the 84 clauses of the Constitution can be read as protecting the right to own slaves without using the word *slavery*—America’s *peculiar institution* (as John C. Calhoun called it). The famous *checks and balances* with power spread across the Senate, House, and Supreme Court, as well as the notions of sovereignty of states—*state’s rights* (and selection of senators by states), limited taxing and other powers of the Federal government, and use of an Electoral College rather than direct voting for the President can all be seen as concessions to the slave owners. (Desmond 2019)

US *exceptionalism* also helps to explain the *low road* taken by the US over its entire history—with less protection for workers and others than in most other nations: the US is dead last in protecting workers from firing among 71 ranked nations; its score for regulating temporary work is tied for second to last; only a tenth of the labor force is unionized; and its richest 1 percent owns 40 percent of the nation’s wealth. (Desmond 2019) Other legacies of slavery include discrimination in housing and schooling, lack of universal health care, and limited coverage for minimum wage laws, unemployment compensation, and Social Security retirement income (that was only gradually increased over time and still discriminates against occupations more likely to employ people of color). While the founding ideals of America are laudable, the reality was very different.

It is also necessary to acknowledge the importance of the appropriation of indigenous land, the genocide of the population, and the exploitation of the natural bounties of the Americas for the development of capitalism in the New World. The primary reason for the extraordinary productivity of labor in the New World was the existence of “free” land for the taking (that is, taking it away from the indigenous population) and the abundance of natural resources.

Just as slaves could be collateralized, the highly productive natural resources of the New World could be collateralized to finance development.

Finally, we need to examine the myth that the rise of capitalism was based on the creation of “free labor.” Graeber (2006) notes that in America in 1600 there were as many white indentured servants as slaves. They largely worked for room and board as well as training, rather than for wages. Those who did work for wages were often paid in tickets or vouchers to be used in the company store or were paid “in-kind” with the wastage of the production process. They might not be paid wages for a year or longer, so they would fall heavily in debt (often to company stores). At that time, the differences among indentured servants, slaves, and wage workers were not great—as Graeber (2006) put it, one could either sell oneself or rent oneself out—in neither case was there much “freedom,” as the owner or the “renter” of the laborer dictated what work was to be done. However, with the rise of racism that proclaimed African slaves to be inherently inferior—in part created to soothe the conscience of Christian owners—the distinction between wage worker and slave grew. Still, Graeber argues that it wasn’t until around 1800 that capitalism began to look like what we would recognize—with relatively “free” labor (in comparison to the situation of Black slaves) paid money wages. This is hundreds of years later than the dates usually assigned to the beginnings of capitalism and free labor.

Slavery and class society, generally, long pre-existed capitalism. (Indeed, even tribal society had slaves, usually captives taken in war.) And the later Greek and Roman societies were substantially based on slavery. These had money and markets, including markets where slaves were bought and sold. Since the time of Mesopotamian society, long-distance trade was conducted on the basis of money with money profits as a goal. And as discussed earlier, throughout the Middle Ages, wars were fought with hired mercenaries and markets were organized for provisioning them. Further, money in the form of credit has been common for thousands of years. So it is true that many of the institutions that are important to the development of capitalism preceded it. Still, capitalism deviates substantially from earlier economic systems that *used* money. Capitalism is a system *based on* money—a monetary production economy.

In sum, the alternative view advanced by Williams almost a century ago and picked up more recently by Graeber (and aired as part of the *NYTimes* 1619 project on the 400th anniversary of the arrival of the first African slaves to America) paints a different picture of the origins of capitalism, defined as a system of M-C-M’ by Marx, or as a monetary production economy by Keynes. Slave owners financed their purchase of slaves captured in Africa by issuing debt collateralized by the slaves themselves—providing the money needed to start the production process. They then produced sugar and cotton (largely for export). Production took place mostly on huge, managed plantations—with

the work done by African and Native American slaves, on land taken from indigenous Americans. Much of the harvest was first processed on the plantations and then shipped to the Old World for further processing. The exports provided the “more money” required to service the debt and provide profits for both the New World plantations as well as the Old World’s shipping and factory businesses. The distinguishing characteristics of modern capitalism can be found in the development of the New World.

In the next section, we closely examine the role played by finance in the modern capitalist economy.

THE ROLE OF FINANCE IN THE CAPITALIST SYSTEM

The distinction between a capitalist economy and a non-capitalist economy was recognized by Keynes, who argued that in a capitalist economy, the object of production is to “result in a larger net sum of money at the end of the accounting period.” (Keynes 1979, p. 66) He defined a barter economy “as one in which the factors of production are rewarded by dividing up in agreed proportions the actual output of their co-operative efforts.” (ibid., p. 66) The barter economy does not preclude the use of money “for transitory convenience,” so “it might perhaps be better to call it a real-wage economy, or a co-operative economy as distinct from an entrepreneur economy.” (ibid., p. 67) In the entrepreneur economy, however, money becomes *essential* and not merely a *useful lubricant* for exchange. Production is a continuing process, so the capitalist develops a relation with a banker who will provide credit to fund working capital expenses (including the wage bill, raw materials costs, and trade credit granted to buyers). The short-term credit is continually rolled over to fund working capital expenses. This type of credit is typically advanced by the commercial banker.

The overriding social purpose of credit money is to provide purchasing power to the capitalist to buy the goods and services needed “today” to produce the goods and services that will be sold “tomorrow.” If this sale takes place at the expected price, the capitalist realizes sufficient money profit to repay the debt and retire the credit money. (Schumpeter 1934)

In the early stages of development (before the late 19th century when railroads and steel manufacturing generated the need for a large stock of expensive and long-lived fixed capital), working capital expenses accounted for the majority of expenses, while investment in fixed capital was a small percentage of national income. During the early period, short-term credit advanced by banks freed the capitalist’s profits for the finance of investment and accumulation of (fixed) capital. Thus, credit in early capitalism can be characterized as financing working capital expenses, which allows production to proceed so that profits can be generated to finance capital accumulation. (Cameron

1967, pp. 36–9, 52) However, as discussed above, the form that capitalism took in the Americas required a huge investment in slavery—not only in purchasing slaves but also in the associated costs: ships and navies, as well as police and instruments of subjugation. Financial markets—New York’s “Wall Street” and London’s “Exchange Alley”—created mortgages and other instruments to finance these considerable expenses incurred both privately and publicly. (In addition, we do not want to deny that credit played a role in pre-capitalist economies.)

In the process of financing working capital expenses, income is generated and a market is created for the goods that will be produced. For example, a portion of the working capital expenditures will be received by workers, who will spend their income on consumption goods. Spending on other working capital (inputs produced by other capitalists) generates more income that goes to workers and capitalists outside the consumption goods sector—and expenditure of that income generates profits, as shown in Chapter 9. The accumulation of physical, productive capital, which increases the capacity to produce goods and services, requires a growing market—otherwise, produced goods cannot be sold at a price sufficient to meet payments on short-term credits, to generate profits, and to finance capital accumulation in the form of money. This, in turn, means that the revolving fund of short-term credits must continually expand.

As argued earlier, if we ignore the government and foreign sectors, it is the production of investment goods (including both intermediate goods used in production as well as physical capital goods) and goods that will be consumed by capitalists that generate aggregate profits. Short-term credit allows rising expenditures on working capital used to produce such goods. Thus, growth of short-term credit generates aggregate profits that fund investment and accumulation of capital. At an individual level, the capitalist borrows a sum of money to produce goods to sell for a greater sum. In the aggregate, this will be possible only if investment goods and capitalist consumption goods are being produced. However, investment raises productive capacity, so it can be undertaken only if market growth is expected. That is, because profits require investment, and because investment requires the expectation of growth, the logic of the capitalist system rests upon growth and capital accumulation. As Minsky put it, investment is forthcoming only if investment is expected in the future—so there will be no profits “today” unless greater profits are expected “tomorrow.”

Heilbroner (1985) argues that the drive to amass capital is the single most important element of capitalism. The form in which the social surplus is accumulated under capitalism is profit, accumulated by the owners of capital. Accumulation of wealth by capitalists is not an end in itself but is the means to further accumulation. In Marx’s scheme, money enters the production process as capital to generate more capital in money form ($m-c-m'$). This continual “expansive metamorphosis of capital” is the essential logic of capitalism.

(Heilbroner 1985, p. 36) A system that relies on growth and short-term credit to fund working capital expenses cannot rely upon commodity money. If credit were to be constrained by the quantity of commodity money available, there would be no guarantee that this quantity would expand to meet the needs of an economic system that requires growth. This is why the capitalist system requires development and extension of the credit system.

Still, the logic of a capitalist system also requires that purchasing power is not extended willy-nilly to anyone who desires credit. Ideally, credit would be extended to capitalists who will enhance productive capacity to increase the output of those goods that can be sold at a price sufficient to cover expenses and leave profits to fund money capital accumulation. Credit creation can be used to funnel resources into the production of investment goods rather than consumption goods to generate a surplus that can be accumulated. (Cameron 1967, p. 292)

However, there is no mechanism in a capitalist system to *guarantee* this result. The individual profit-seeking behavior of those who advance credit, however, ensures that loans will normally only be made to those who are highly likely to repay them (with interest). Furthermore, individual firms are forced to produce commodities for sale to obtain money, since they cannot produce money but are forced to borrow it. Thus, profit-seeking behavior necessarily rations credit, and repayment of credit necessitates orientation toward the market. Traditionally, financial institutions try to develop close personal relations with borrowers and then apply various rules of thumb (such as collateral requirements, market share of the borrower, credit history, and so on) when deciding whether to make a loan. The financial institution operates as the expert judge in determining who is likely to repay loans, or, as Schumpeter put it, the banker is the *ephor* of capitalism. (Schumpeter 1934) Innovations and successful experimentation have gradually expanded the types of borrowers deemed creditworthy—again, there is no guarantee that credit rationing will necessarily generate the right amount of credit and place it in the correct hands. Thus, failure and bankruptcy are a necessary, if normally infrequent, outcome of a privately motivated capitalist system.

As fixed capital became increasingly important with the development of the industrial system and as capital became more complex and expensive, investment could not easily be funded out of retained profits. Gradually, firms shifted from “internal” finance of investment (retained profits) to “external” finance (borrowing), in which short-term credits would continue to be used to fund production of the investment goods, and long-term debt would be used to fund a position in the finished capital goods. In the aggregate, it is the short-term credit (which finances working capital expenses) that generates the income used to purchase the long-term debt, which in turn funds the position in the capital asset.

In orthodox thought, the original purpose of the development of banking is to intermediate between savers and investors—thus supplying the funds for long-term investment. However, the provision of long-term credit through bank “intermediation” is necessarily a late development, coming after the public has become accustomed to holding and using the short-term liabilities of banks—that is, after the “banking habit” has become well developed. As Bagehot (1927) argues, and as historical evidence generally shows, banks begin by issuing notes, and only later do deposits become important, after the public has experience with “bank intermediation.” At this stage, the financial *system* can finally mobilize capital by issuing deposits as a short-term liability and by making long-term loans to enable firms to finance capital purchases. (Note that this does not imply that an individual financial *institution* must provide both short-term deposits and make long-term loans, as commercial banks typically provide the deposits while investment banks arrange long-term funding of investment—see below.)

Thus, the historical order of events shows that orthodoxy clearly has reversed the process through which investment is funded. Banks do not begin as intermediaries that accept deposits of “savers” and then make loans to “investors,” for this would assume that the public has already developed the banking habit. This habit is the result of public experience with short-term bank liabilities, which have been created as banks extend short-term credit to finance working capital expenses. (Cameron 1967, p. 295) After the public becomes accustomed to receiving banknotes as payment for wages and to using them for consumption expenditures, it will gradually accept bank deposits as a form of intermediation.

In practice, individual banks typically try to match the maturity of assets and liabilities. The banking business becomes segmented, with commercial banks making loans to finance working capital expenses and issuing notes and then deposits, while investment banks handle placements of longer-term bonds and equity shares. The production of investment goods is financed by commercial banks, but the final sales of expensive and long-lived capital goods are funded by investment banks that place the savings generated from the production process into longer-term financial assets (corporate bonds and stocks).

Just as the logic of a capitalist system requires that it cannot be constrained by commodity money, its logic also requires that it cannot be constrained by saving. The production of an investment good is financed through short-term credits that generate profit income. In the absence of long-term credit, positions in investment goods are financed by this profit income directly. However, the direction of causation is clear: it is the spending on investment goods—financed through short-term credits—that generates the profit income that can be saved in the form of final sales of investment goods. The transition to the external finance of positions in investment goods does not change the logic

of this system: it is still the production of the investment good that generates profit income. However, in this case, profits need not directly fund sales of finished investment goods since long-term financial assets can be sold to those receiving profit shares. Alternatively, surplus income can be used to buy short-term financial assets issued by investment banks or other financial institutions (such as pension funds, university endowments, and sovereign wealth funds), which take up the long-term instruments.

CAPITALISM AND THE DEVELOPMENT OF A MONO-RESERVE SYSTEM

In a well-functioning capitalist economy, most money will take the form of privately supplied credit, as short-term loans fund working capital expenses based on note issues or, later, demand deposits. In the absence of government backstops, the value of the credit is maintained by the value of the loans. Frequent default and subsequent failure of banks would make the public wary of them and impede the development of the banking habit. To placate the public, bank liabilities would often (but not always) be made convertible into some other asset that would be held in reserve to meet the demand for redemption. Development of the banking habit and normal operation of the financial system require that bank liabilities are rarely redeemed for the reserve asset. At first, the reserve asset might consist of gold coins, to which the public had already become accustomed. However, a credit system based on a reserve of gold coins is unworkable—the reserves of an individual bank can never be sufficient to meet a run on a bank. When banks were small and regionalized, a bank run would be stopped not by its reserves, but by guarantees of credit-worthiness provided by trusted individuals. A reserve would suffice only if the public did not exercise its right to redeem bank liabilities.

A financial system with large, national banks could not rely on individual gold reserves. The first step toward an integrated banking system required an integrated system of reserves. As discussed in Chapter 3, in England, this took the form of a system of country bankers that relied on London banks to provide reserves as needed. While the London banks could mobilize reserves to meet a limited local run on reserves, they were still helpless to stop a general run. The development of the Bank of England as the mono-supplier of reserves could potentially provide sufficient reserves (Bank of England notes) to stop any run. However, if it viewed itself as a private bank, it would not be willing to do so. Only a central bank that is willing to issue its liabilities without taking a narrow view of its own profitability can fulfill the lender of last resort function. As Bagehot insisted, the lender of last resort recognizes that it must operate against its self-interest to save the system—which the Bank of England was willing to do by the second half of the 19th century.

The logic of the capitalist system requires that private credit normally functions as money and that reversion to reserve money or metallic coins is an irrational aberration and a rare occurrence. A run out of credit money causes debt repudiation and destroys the value of accumulated capital assets. Orthodoxy would have us believe that the normal functioning of a capitalist economy requires that spending is related to M1, defined as currency and demand deposits, where the quantity of demand deposits is strictly constrained by reserves provided by deposits of currency or by the central bank. According to orthodoxy, banks learned that they needed to hold only a fraction of a deposit on reserve and could loan the rest. That is, the logic of the orthodox system can be stated as deposits make reserves, which make loans.

The extensive use of demand deposits, and the relation between demand deposits and reserves, are developments that came after the public had become accustomed to the use of a wide variety of private liabilities. Banks first operated by loaning their own liabilities (notes) that were gradually accepted by the public and later deposited in banks. The logic of a capitalist system is that loans “make” deposits—as the Horizontalists put it—because the system cannot be constrained by deposits. The existing system represents the result of the evolution of the financial system away from a wide variety of privately issued credit instruments to a system based on demand deposits backed by the liabilities of a central bank. All of this is consistent with the endogenous money approach presented in Chapter 6.

Indeed, it is instructive to note a steady decline in the velocity of M1 (national income divided by M1—bank notes plus deposits plus currency) in England from between 4.0 to 5.2 over the period 1688–1775 to just 1.6 by the last quarter of the 19th century. (Cameron 1967, p. 42) This clearly indicates a transition from a system with a wide variety of credit instruments to a system based on deposits of commercial banks and contradicts the orthodox story in which banks gradually discover the existence of a deposit multiplier. Credit money is the normal case, and current institutional relations, which may obscure this fact, are the result of the evolution of the financial system.

THE VALUE OF MONEY

In orthodoxy, the value of money is the inverse of the price level—if prices rise, the value of money falls. Prices, in turn, are determined by supply and demand. At the micro level, prices reflect relative scarcity and efficiently allocate the limited supply among competing demands. In the aggregate, the price level is determined by the demand and supply of money. Increasing the money supply raises the price level and reduces the purchasing power of the money unit. In most orthodox approaches, money is demanded for its usefulness as a medium of exchange—not because it provides utility. Money is neutral in the

long run, although it can have real effects in the short run due to various kinds of fooling or illusion. Money has a functional value, but in and by itself, it has no value unless it takes the form of coined precious metal—in which case it is the metal's scarcity and value in alternative uses that give money value.

For Monetarists, the relationship between money and the price level is direct: *inflation is always and everywhere a monetary phenomenon*. Too much money causes aggregate demand to outstrip aggregate supply, prices to rise, and the value of money to fall. For orthodox Keynesians, the link is less direct. An increase in the money supply lowers interest rates, increasing interest-sensitive spending. The main cause of inflation is excessive demand for labor and, as the unemployment rate falls, wage pressures increase costs while higher income boosts consumption, causing the price level to rise. In recent years, the link has become less direct as the New Monetary Consensus took hold: inflation is largely caused by expected inflation. In the short run, monetary policy plays a role, both because of sticky wages, interest rates, and prices and by providing signals that influence expectations of inflation. Still, money's role is largely reduced to short-run effects and its value is determined by purchasing power.

In all three of the main heterodox traditions—Marxist, Institutional, and Post Keynesian—money plays an important, non-neutral role in the economy. Each emphasizes the monetary theory of production—production begins with money to purchase the needed inputs, with the expectation of generating a greater quantity of money at the end of the production process. In heterodox approaches, micro-level prices are largely determined on the supply side by production costs. Especially in the Post Keynesian literature, money is endogenous and its growth is linked to the wage bill: banks make loans and create deposits to finance wages. However, wages are not the only costs, so loans are also necessary to cover other production costs. Short-term loans are rolled over each production period and increase if production is increasing. If wages and other costs are rising, more money will likely be created—but it would be overly simplistic to say that “money causes inflation” as money's growth is an effect, not a cause of rising wages and other costs.

Many heterodox economists take a mark-up approach to pricing: firms add a mark-up over costs to ensure profits. There are a variety of views about what to include in costs and about the nature of the mark-up—but we will not pursue the details here. In short, prices could rise either because costs are rising (largely supply-side driven inflation) or because mark-ups are increasing (which could be due to rising pricing power on the supply side, or because high demand enables price-setters to raise prices; both of these probably played a role in the high inflation following the COVID recession). Finally, prices at the micro level can be pushed up by higher demand facing inelastic supply. As Keynes (1964, Chapter 21) argued, the elasticity of supply differs across

different kinds of output, so the response of prices to an increase in aggregate spending depends on where it is directed.

Is the value of money simply the inverse of the price level? While heterodoxy's explanation of the causes of inflation is different from orthodox views, heterodox traditions have not generally delved into the determination of the *value* of money. There are, however, two notable traditions that do address the question of money's value: Keynes's liquidity preference theory interpreted as a theory of value for assets, and Marx's labor theory of value applied to money. We will begin with the labor theory of value and then address the liquidity preference theory of value, but first, we need to discuss what we mean by value theory.

1. What is Value Theory For?

Most economists today focus on price and either ignore or put less emphasis on value as a subjective notion. But economics began as moral philosophy, and earlier economists sought to do more than provide a theory of price. The goal, as Graeber put it:

... has been to understand the workings of any system of exchange (including free-market capitalism) as part of larger systems of meaning, one containing conceptions of what the cosmos is ultimately about and what is worth pursuing in it. Such systems of meaning meant that the kind of moral and ethical questions that Aquinas or Smith felt were at the heart of the matter could not simply be pushed aside. (Graeber 2005, p. 443)

Or, as Heilbroner argued,

The general problematic of value, as I see it, is the effort to tie the surface phenomena of economic life to some inner structure or order. This problematic arises because economics is unavoidably involved in two intimately related but essentially distinct tasks. One of these is the investigation of various empirical aspects of the process of social provisioning ... Equally necessary for the existence of what we call economic thought is a level of abstract inquiry—an inquiry directed not at the “facts” of economic life but at some structure or principle “behind” the facts ... It necessarily looks beyond appearances for essences... Economics now becomes an inquiry into the systemic properties, the structural attributes, the tendencies and sometimes even the *telos* of the provisioning process. (Heilbroner 1988, pp. 105–106; emphasis in original)

Economists have variously sought *laws* of the behavior of individuals (utility maximization; rationality) or of the system (*laws of motion*: forces that lead to a general equilibrium). Given the importance of prices in a capitalist system, they have been the focus of economists of all traditions. “Value theory

is therefore indispensable for understanding how the capitalist system, largely guided by price stimuli, tends toward some kind of determinant outcome.” (op. cit., p. 107) Heilbroner quotes an insightful thought experiment posed by Adolph Lowe: “Suppose that a universal amnesia were to wipe out the knowledge of all present prices, would there be a rule for reestablishing them?” (op. cit., p. 108) Most economists would presumably answer in the affirmative—prices are not randomly determined. Our theory of value informs our beliefs about how the *deep structure* of the economic system would restore a system of prices.

In orthodoxy, money plays no important role in those processes. Prices are relative—measures of scarcity. However, in most heterodoxies, money plays a role beyond that of medium of exchange—although exactly what role(s) it plays varies across traditions. For at least some, money is part of that deep structure of the capitalist system. As discussed above, Heilbroner (1985; see also Wray 1990) explains that money is central to the *internal logic* of the capitalist system: the drive to amass capital in money form is the single most important element of the system. As discussed above, it is what makes capitalism truly different from other social organizations, and to a great extent makes it possible to examine capitalism as an *economic system* that is somewhat dis-embedded from the social system as a whole—operating according to a *logic* that is capable of economic analysis.

Graeber nicely contrasts the heterodox approach against that of orthodoxy:

The first [heterodox] begins by imagining some total system or structure—a society, a world-system—and then try to understand how it is maintained and reproduced over time. The other [orthodox] starts with individual actors pursuing something, and sees society largely as the effect of their actions (here economics and its derivatives, like rational-choice theory, have been the paradigm). (Graeber 2005, p. 445)

In sum, for heterodoxy, money plays the key role in the structure of the capitalist system that gives it a logic capable of being studied as an *economic system*. We will turn to two heterodox theories of value and of the role of money in that deep structure and argue that these are not alternatives but rather address different parts of that structure. The first is the labor theory of value associated most closely with Marx, and the second is the liquidity preference theory of value derived from Keynes’s liquidity preference theory of asset pricing. As there are two price systems that play critical roles in the deep structure, we need both theories of value. We will argue that the labor theory of value helps us to understand the pricing of output, while the liquidity preference theory of value sheds light on asset prices. The two come together in the market for physical capital because the investment decision involves both.

2. A Labor Theory of Value Applied to Money

There are two competing views of Marx's approach to the value of money. (Foley 1983) The first interprets Marx as arguing that the labor hours required to produce the commodity money determine the value of money—hence money is just like any other commodity produced by labor. While there is evidence to support this interpretation, it conflicts with other discussions by Marx. Further, with the abandonment of gold standards (and the limited use of them throughout money's history), this view cannot be applied to today's so-called *fiat money* sovereign currency systems. Finally, the gold-commodity view of money is not consistent with either the history of capitalism or with its underlying structure: money cannot be something produced by labor—it must be external to the system of commodity production. (As Keynes implied, if money grew on trees, labor easily could be “employed” to harvest it—resolving the unemployment problem.)

For those reasons, we will focus on the second interpretation of Marx: money's value is determined by socially necessary labor time—not at the level of production of any particular commodity but rather at the aggregate level. Nor is money's value determined (in an inverse way) by the price of output. Labor value does not equal price, and indeed, labor values must deviate from prices (the so-called transformation “problem”—as we'll see, that is a feature, not a bug of capitalism).⁴ (Foley 1982, 2018)

As discussed in the *introduction* of this book, Levine argues that production for market exchange already presumes specialization and separation of producer from consumer. This would have been too risky in the absence of markets and money. Production for market is “commodity production” and requires a system of commodity relations with an external measure of value. While a single commodity can satisfy an individual need, for relative values, we need to measure in terms of something that is not a commodity—we need a universal measure of value. Money is the external measure of value and has no “value” itself in terms of satisfying *individual needs*. In Levine's view, money *is* value, “emancipated” from individual needs.

Similarly, Foley (1983, 2018) argues that money is an expression of abstract labor. It is not derived from a commodity—money is the unit in which promises to pay are measured. (This is also the MMT view, as discussed next.) Labor creates all value, but only money values are observed as exchanges occur at money prices. Labor values are preserved in exchange as money prices redistribute value to equalize money profit rates across processes with different organic compositions of capital (that is, ratios of “living” labor used in production to “dead” labor embodied in capital).⁵ Money values—not the unobserved labor values—drive production decisions where the goal is to end up with more money later. At the aggregate level, money value measures the

aggregate of labor value. And, money value of profits is the only measure of success (from the point of view of capitalists that organize production—they want money, not labor value).

Foley presents the “labor theory of value as the claim that the money value of the whole mass of net production of commodities expresses the expenditure of the total social labor in a commodity-producing economy ...” (1982, p. 37) “A unit of money, in this approach, can be thought of as a claim to a certain amount of the abstract social labor expended in the economy”. (ibid.) However, money prices deviate from labor values for individual commodities, but in a predictable manner (according to the organic composition of capital). (ibid.)

Foley follows Marx (and Keynes 1964, Chapter 4) in adjusting measured labor time to correct for “differences in the intensity of work, the skill of workers, and the relation of the technique of production to the current social standard”—that is, he uses “simple, abstract, socially-necessary labor.” (Foley 1982, pp. 38–9) The value of money is then defined as the “ratio of aggregate labor time to aggregate money value added.” (op. cit., p. 41; see Chapter 11 for a numerical example.) The level of the money wage is the cost of the socially determined standard for the subsistence of workers. At the aggregate level, “profits are proportional to aggregate unpaid labor time,” although individual prices can—and generally must—deviate from labor value to redistribute profits. (Foley 1982, p. 42)

Money plays a role not only in distributing surplus value but also in exploiting labor—as workers and capitalists struggle over the level of money wages. The wage is a claim on a share of the labor embodied in commodities, and workers bargain over a money wage, not a bundle of commodities. Following the Smithian and classical school’s long-period approach, mobility of capital and labor ensures a tendency for market prices to move toward *natural prices* that will equalize the rate of profit on capital and the rate of exploitation of workers across employments. The profit rate is determined as net money revenues relative to money capital invested and is a portion of the surplus value that is realized in money form and created by unpaid labor power.

This is extracted at the aggregate level and forms a social fund (“capital”) that is distributed among individual capitals through competition. The aggregate money surplus (what Minsky 1992 called gross capital income) also supports other incomes such as rent, royalties, commercial activities that do not generate surplus, and the financial sector (interest, capital gains).⁶ It is the money price system that distributes the surplus among claimants. This presupposes a monetary production economy, for otherwise there would be no way to distribute the surplus to equalize profit rates. The price system provides the signals that mobilize capital movements to ensure this tendency. Unlike the view of neoclassical economics, prices are not determined to equalize “supply

and demand” but to distribute a surplus in money form that will tend to equalize the rate of profit.

Similarly, money wages mobilize labor so that the rate of exploitation, which is the ratio of the money value of the surplus to the money wage, tends to equality across employments—rather than to equate “supply and demand” in a labor “market”. (Foley 2018)⁷ While the wage is proportional to effort, it only rewards a part of the effort. The natural (long period) price of commodities is also proportional to labor effort: the *monetary equivalent of labor effort*. That ratio is a measure of the degree of exploitation, which tends to equality across lines of production as commodity prices tend to the natural prices. Hence, the same competitive processes that redistribute surplus value to equalize the rate of profit also tend to equalize the rate of exploitation. Labor effort and labor values are unobservable but underlie the money wages, money prices, and money profits that are the bases of decisions.

In the Marxian approach, all the wages paid represent a claim on subsistence commodities. Since the aggregate money value of subsistence commodities is equal to aggregate wages paid to produce them, plus nonwage costs and the money value of surplus labor, wages from producing other commodities (i.e., commodities for investment, government, and exports) are also spent on subsistence commodities. And because the sum of the wage bill across all sectors only equals paid labor, the unpaid labor surplus cannot be realized without additional spending—by capitalists, rentiers, unproductive labor, government, and sales to foreigners.

3. Keynes’s Liquidity Preference Theory of Value

This leaves us with Keynes’s approach to liquidity preference and interest rate determination discussed above in Chapter 7. Townshend (1937) recognized that Keynes presented a general theory of asset pricing, or what we could call a liquidity preference theory of value.

The “fundamentalist” Post Keynesians—such as Davidson, Minsky, and Kregel—emphasize the approach taken to money in Chapter 17 of the *General Theory* that interprets liquidity preference as a theory of asset prices.⁸ Minsky, in particular, focuses on the “two price systems”: one for current output and the other for assets. These come together in the demand price and supply price for newly produced capital assets. Newly produced investment goods have a supply price determined in the current output price system, plus financing costs (that cover lender’s risk). The demand price for capital assets comes out of the asset price system as productive capital competes with financial assets. The expected return to owning and operating the capital assets plus borrower’s risk must be high enough relative to financial assets or investment goods will not be produced. This is essential to Keynes’s theory of effective demand—as

investment drives demand. Here, money's importance is paramount for two reasons: its total return (mostly due to liquidity) sets the standard for all other assets, and because the terms on which access to external finance is available affect the level of investment.

To recap and deepen the discussion in Chapter 7, money is the most liquid asset whose return is determined by its liquidity less its carrying cost, which for money is greater than that of any other asset. Any asset that can be held through time has a total return equal to its yield, minus the carrying cost, plus its liquidity premium. From Sraffa (1932), Keynes borrowed the concept of the own rate of an asset that is determined by the relationship between its spot price and its forward price. In Chapter 17, he gives an example using wheat: if 100 units of wheat delivered today ("spot") have the same exchange value as 105 units delivered a year from now ("forward"), then the wheat rate of own interest is 5 percent.⁹ Producible assets (i.e., investment commodities, although anything that can be held through time has an own rate) are produced until their own rate falls in line with the own rate on money. In this respect, money "rules the roost." The own rate on money is simply the base interest rate set by policy—the fed funds rate in the USA.

If spot prices equal forward prices, there is no incentive to increase production; if the spot price is above (below) the forward price, then there is an incentive to produce more (less). This is equivalent to a situation in which the current demand price for the asset exceeds its supply price.¹⁰ As production of each type of producible asset (capital or other commodity) increases its supply, the spot price tends to fall¹¹ relative to the forward price. Aggregate production proceeds until all own rates fall in line with the own rate on money—"the" interest rate. (See Kregel 1988.)

Sraffa had argued that if "loans were made in terms of all sorts of commodities, there would be a single rate which satisfies the conditions of equilibrium, but there might be at any one moment as many *natural* rates of interest as there are commodities, though they would not be *equilibrium* rates." (Sraffa 1932, p. 49) Chapter 17 of the GT follows the argument made earlier by Sraffa that this natural rate on money and the level of effective demand is associated with the equalization of the natural rates of all other assets: "In equilibrium the spot and forward price coincide, for cotton as for any other commodity; and all the *natural* or commodity rates are equal to one another, and to the money rate." (Sraffa 1932, p. 50)

However, in the GT, Keynes distinguishes between the natural rate and the neutral rate of interest on money. There is a different natural rate associated with each point of effective demand, but only one of these is *neutral*—the one consistent with the level of effective demand associated with full employment. At each natural rate, entrepreneurs are employing the quantity of labor required to produce the amount of output they believe to be warranted by their

expectations of profits—but this need not be at full employment. Indeed, due to the special properties of money, the point of effective demand tends to be reached at a natural rate below full employment.

Keynes rejected the notion that money determines the price of output, putting it this way:

Money, and the quantity of money, are not direct influences at this stage of the proceedings. They have done their work at an earlier stage of the analysis. The quantity of money determines the supply of liquid resources, and hence the rate of interest, and in conjunction with other factors (particularly that of confidence) the inducement to invest, which in turn fixes the equilibrium level of incomes, output and employment and (at each stage in conjunction with other factors) the price level as a whole through the influences of supply and demand so established. (JM Keynes in the *Introduction* to the French edition of *The General Theory of Employment Interest and Money*, 1939)

Instead of determining prices after production has taken place (that is, in the sphere of circulation), money does its job earlier in the sequence—in the investment decision, not in the pricing of final output. Note that this opposes the mainstream monetarist view that the quantity of money determines the aggregate price level. Instead, Keynes argued that “The general price-level depends partly on the rate of remuneration of the factors of production which enter into marginal cost and partly on the scale of output as a whole, i.e. (taking equipment and technique as given) on the volume of employment.” (Keynes 1964, p. 194)

Importantly, prices do not move to clear markets. Equilibrium at natural prices is consistent with the point of effective demand, not with full employment.

4. Integrating Two Theories of Value

While Keynes does not explicitly adopt the labor theory of value, his method of analysis is similar. He adopts only two fundamental units of quantity, “quantities of money-value and quantities of employment,” and he adopts two measuring units: the wage unit and the quantity of labor hours.¹² In his view, this is the only way to get around problems of measuring the volume of heterogeneous output and the general level of prices. (Keynes 1964, p. 43) Aggregate supply is then a function of employment, and the “ordinary supply curve” depends on employment and the expected user cost associated with that level of employment.

Output is measured by aggregating the employment involved in producing each commodity—since heterogeneous output cannot be aggregated. Its value is measured as “the number of hours of labor paid for ... on the existing

capital equipment, hours of skilled labor being weighted in proportion to their remuneration.” (op. cit., p. 44) This avoids the “well-known, but unavoidable, element of vagueness which admittedly attends the concept of the general price-level ...” (Keynes op. cit., p. 39) Here Keynes was again following Sraffa’s lead in rejecting any attempt to construct a price level for output as a whole that is linked to the money supply.¹³

In the “fundamentalist” approach to Keynes, the importance of money is not found in its relation to an overall “price level” (that is necessarily a chosen index from among a limitless number of possible indices) but rather in money’s impact on the *relative* money prices of commodities and hence on the level of effective demand, which is a function of money’s own rate. Its “spot” value in terms of commodities is its purchasing power of labor—the wage unit. The relative wage unit is that for an hour of ordinary labor effort. As in the Marxist approach, the *value of money* is not just the flip side of the aggregate price level. Price indices are arbitrary and heterogeneous output cannot be simply added up. We use only two measures of value, labor hours and the money wage unit, *because there is no alternative*. The value of money is equal to the amount of labor time it can purchase. The *price* of output is complexly determined and must be higher relative to the wages paid to produce it to cover finance and other overhead costs (advertising, management, taxes) associated with the version of capitalism in which the output is produced. At the aggregate, this *surplus* depends on the share of investment, government spending, and net exports in total output. Gross capital income (profits) must be redistributed through the price system toward firms with higher business costs, higher capital ratios, and more financial debt.

The liquidity preference theory of value supplements the labor theory of value, which applies only to commodities that are produced by labor. Liquidity preference is a value theory for assets that are not produced by labor, with the exception of physical capital. The two theories of value—and the two price systems—meet in the investment decision, which Keynes saw as the economic driver of the capitalist system. The price system for commodities produced by labor redistributes labor value to equalize profit rates; the price system for assets equalizes expected returns on assets. The investment decision requires both as the investor in capital weighs expected returns against the highest expected return on financial assets.

Why do we need two theories of value? The labor theory of value helps us to understand the logic of commodity production and how money prices equalize rates of profit and exploitation across different production processes. The liquidity preference theory of value adds an understanding of the formation of asset prices. In Keynesian theory, the most critical decision is the investment decision—it is the anchor for the theory of effective demand. Investment is the driver of the system not only because of the multiplier effect but also because

investment is the source of aggregate profits. As Minsky put it, investment spending (but also government deficits and a current account surplus) *forces a surplus*. Marx also recognized that the realization of surplus value requires an additional source of demand for consumption goods beyond that generated by the wage bill in the consumption sector.

What liquidity preference theory adds is a theory of the investment decision consistent with a capitalist system in which the purpose of production is to end up with more money. Capitalists will invest in physical capital only if the expected return can compete with the return expected from purchasing financial assets. Keynes saw liquidity preference as creating a hurdle for investment returns: given a preference for liquid assets, capital assets must earn a higher return to compensate for giving up liquidity. He argued that this is what keeps investment too low to generate full employment.

WHAT DOES VALUE THEORY CONTRIBUTE TO OUR UNDERSTANDING OF CAPITALISM?

We have focused on the approaches of Marx and Keynes to value theory, which share several similarities. There are two measuring units: socially necessary labor time (or “ordinary labor” in Keynes’s terminology) and money value (Keynes’s wage unit). Capitalism is a monetary production economy, driven by the expectation of money profits. At the aggregate level, Keynes measures output as the total of paid labor hours—which is precisely Marx’s view as interpreted above. In both approaches, money’s value is equal to the wage paid for an hour of (average) labor. And money is the sole external measure of value.

Marx’s approach emphasizes that the “unique thing about capitalism” is that “it allowed labor to become an abstraction,” turning labor power into a commodity that can be bought and sold. That commodity is the “*capacity to work*.” (Graeber 2005, p. 450)

What makes this possible is the use of a specific symbolic medium of value: money. For Marx, money is a symbol. It represents the *value* or importance of labour. It can do so by incorporating it into a total market system, because for Marx the real value of a product is not (as Ricardo claimed) how many hours of work went into making it, but the proportion of the total amount of labour in the entire economy that went into making it. This proportion can only be determined through the market; that is through the use of money. (Graeber 2005, pp. 450–1)

Thus, value is tied up with money, but not directly with a commodity’s price. While we can count labor hours, these hours are specific to the tasks to which they are employed. The price of a particular commodity produced by labor is not equal to its embodied labor value, as prices must redistribute surplus labor. In answer to the question raised by Adolf Lowe—if we woke up tomorrow

with amnesia about prices, could we reconstruct them?—the answer is yes (albeit with some difficulty). The price system(s) forces a surplus (as Minsky said) and redistributes it to equalize the rates of profit and exploitation. But there are two price systems: one for current output and one for assets. We need two theories of value, and money is the measure of value in both.

Keynes's approach to measuring units—money value and labor hours—is similar to Marx's, as we saw above. However, what is perhaps more important is that he brought values in the sense of morality back into (non-Marxian) economics. Keynes does not take wants as given but rather wishes we might make them desirable wants. As Skidelsky argued, Keynes believed that "[t]o make the world ethically better was the only justifiable purpose of economic striving." (Skidelsky 2009, p. 133) Keynes saw capitalism "as a necessary stage to get societies from poverty to abundance, after which its usefulness would disappear." (ibid., p. 135) Following G.E. Moore, Keynes believed that *good* is objective, that we know what is good, and "that which is to be maximized is not happiness or pleasure, but goodness." (ibid., p. 137) Keynes's view was that the *love of money* is a neurosis to be tolerated only until we achieve the abundance that would allow us to realize the "*Economic Possibilities for Our Grandchildren*." (ibid., p. 144)

As Graeber puts it, in this sense of the term, values are "conceptions of the desirable," "ideas about what people *ought* to want." (Graeber 2005, p. 446) The important point is that both Marx and Keynes saw the immorality of capitalism as a system that pursues profit above all else. In Keynes's theory, the result is unemployment and excessive inequality (also instability, as Minsky argued)—both attributed to the use of money. In Marx, inequality and instability are the result of the exploitation of labor (unpaid labor time). Graeber argues that

Marx did not propose a labour theory of value mainly as a way to explain price fluctuations, but as a way of connecting economic theory with broader moral and philosophical concerns. For Marx, *labour* was more or less identical with human creativity ... The unique thing about capitalism, Marx held, was that it allowed labour to become an abstraction. This was because capitalism turns labour into a commodity, something that can be bought or sold, and what an employer who hires a labourer buys is an abstraction, that labourer's *capacity* to work. What makes this possible is the use of a specific symbolic medium of value: money. (Graeber 2005, p. 450)

Money is "fetishized" as "wage labourers only go to work in order to get money" (Graeber 2005, p. 451) and capitalists only hire them to get *more* money. As Keynes put it, capitalists have no other desire in the world. The capitalist economic system itself *values* the money token, not the ability to satisfy basic human needs.

How can the heterodox approaches that derive from Keynes's thought be synthesized with the Marxian approach? Wages owed are denominated in the state's money of account, as are commodity prices and asset prices. Payment of money clears the wage obligation to workers; money received as wages represents a *claim ticket* on commodities, redeemed for subsistence consumption. If we expand our analysis to include private money, this too takes the form of credits and debits denominated in the money of account. Banks advance money in the form of deposits to pay wages and other costs as firms produce commodities for sale; receipts allow firms to meet their obligations to banks by redeeming bank money. Market prices determine the quantity of money for which commodities can be exchanged at a point in time, but in the long period they approach natural prices that equalize profit rates and rates of exploitation across commodities. There is thus a proportion between labor hours and money, what Foley terms the monetary equivalent of (ordinary) labor time.

While orthodoxy views the state as *intervening* (for better or worse) in a market economy, heterodoxy sees the state as fundamental to the capitalist economy. MMT brings the state into the analysis not only to choose the money of account and issue currency redeemable for obligations to the state, but also to issue the ultimate money for the clearing of private monetary obligations (i.e., central bank reserves). We can also bring in *political economy* aspects of the state's role in primitive accumulation, division of labor, wage labor, and markets. (Forstater 2004, 2005)¹⁴ This helps to resolve problems with the usual approach to money identified by Levine (1983): imagining a pre-existing market economy based on commodity exchange but without money.

Finally, as discussed above, the value of labor is not determined at the individual level but at the level of the economy as a whole: "the proportion of the total amount of labor in the entire economy that went into making it." (Graeber 2005, p. 451) This is the difference between abstract labor (economy as a whole) and concrete labor (hunting deer)—commodities do not exchange on the basis of concrete labor hours (as in Smith's example where he argues that if it takes twice as much labor to kill a beaver, then one beaver will trade for two deer) but rather are sold for money that represents a claim ticket on abstract labor power.

MMT insists that we should have *the state* in the picture from the very beginning: it creates a money of account, imposes obligations, and issues the currency that can be used to pay those obligations. In this way, we do not have to imagine the operation of a capitalist economy without a state and then add the state and its currency to a pre-existing, stateless system. There is no capitalism without a state that chooses the abstract measure of value—the money of account. There is no capitalism without a state that provides a legal framework, enforces contracts, and provides the institutional framework within which capitalism functions. This does not mean that the state dominates or

operates separately from capitalist relations, but rather that it plays a socio-political role that goes beyond “spending and taxing.”

In the next chapter, we will examine MMT’s best-known policy recommendation: the pursuit of full employment through a Job Guarantee program. We will see how this can not only achieve and maintain true full employment but can also anchor the value of the currency in terms of the money value of the wage unit.

NOTES

1. For research on the role played by New World slavery in the development of capitalism, see: Eric Williams (1944), *Capitalism and Slavery*, Chapel Hill: University of North Carolina Press; David Graeber (2006), “Turning Modes of Production Inside Out: Or, Why Capitalism is a Transformation of Slavery”, *Critique of Anthropology*, 26(1), pp. 61–85; Seth Rockman (2014), ‘Review Essay: What Makes the History of Capitalism Newsworthy?’, *Journal of the Early Republic*, 34(Fall), pp. 439–66; P.R. Lockhart (2019), “How slavery became America’s first big business”, Vox, <https://www.vox.com/identities/2019/8/16/20806069/slavery-economy-capitalism-violence-cotton-edward-baptist#>; and a number of articles in the *New York Times Magazine*: <https://www.nytimes.com/interactive/2019/08/14/magazine/1619-america-slavery.html> from its 1619 project.
2. See Jeffrey Robinson, *Who We Are*, 2022 Documentary: <https://www.sonypictures.com/movies/whoweareachronicleofracisminamerica>; discussion here: https://www.democracynow.org/2022/1/14/who_we_are_film_jeffrey_robinson.
3. While the Netherlands dominated the early Atlantic shipping of slaves, they were soon replaced by Portugal, which was then surpassed by Britain: “The British and Portuguese account for seven out of every ten transatlantic slaving voyages and carried nearly three-quarters of all people embarking from Africa destined for slavery” (David Eltis (2002), ‘The Volume and Structure of the Transatlantic Slave Trade: A Reassessment’, *The William and Mary Quarterly*, 58(1), pp. 17–46, (Electronic Document) January 2002). <https://www.nps.gov/ethnography/aah/aaheritage/histcontextsd.htm#:~:text=Where%20did%20enslaved%20Africans%20come,Americas%20than%20any%20other%20region>.
4. As Foley (1982) argues, while it is possible to come up with the assumptions required to transform labor values directly into prices, in the general case labor values must deviate from prices. The most important reason is that prices must redistribute surplus value to equalize profit rates on capital (as a trend, not necessarily at any given point in time), which causes prices to deviate from values because labor-capital ratios vary across firms. In addition, market pricing power, as well as what Minsky called “business style,” varies across firms and sectors. See below.

5. Only living labor produces surplus value; the money price system redistributes surplus value to processes with high capital-to-labor ratios ("dead labor-to-living labor ratios"). Firms with high capital ratios will need higher money prices relative to the value of labor embodied in the commodities produced. Note the similarity to Minsky's argument: high capital ratios require higher profits to validate the investment.
6. Owners of scarce resources can obtain monopoly rents. See Minsky (1992).
7. This is an updated version of the *New Interpretation* of Marx's theory of value. See Foley (2018).
8. This can be presented as a theory of value for asset prices—that is, for anything that can be held through time. See Townshend (1937), and Wray (1990) and (1992).
9. We can also calculate wheat's own rate in terms of money, but this requires knowing money's own rate—the interest rate on money. If that is 5 percent, and if 100 units of wheat delivered spot costs \$100 while the forward price is \$107, then selling wheat for \$100 and earning 5 percent produces \$105 over the year, and that would purchase only 98 units of wheat for forward delivery. The wheat-rate of interest is then minus 2 percent in terms of money. (Keynes 1964, p. 223)
10. As Sraffa put it, "if, for any reason, the supply and the demand for a commodity are not in equilibrium (i.e. its market price exceeds or falls short of its cost of production), its spot and forward prices diverge, and the 'natural' rate of interest on that commodity diverge from the 'natural' rates on other commodities ... It will be noticed that, under free competition, this divergence of rates is as essential to the effecting of the transition as is the divergence of prices from the costs of production; it is, in fact, another aspect of the same thing." (Sraffa 1932, p. 50)
11. With generally constant costs, raising output lowers markups, and prices fall along with quasi-rents.
12. Keynes (1964), p. 41. More specifically, he uses "ordinary labour as our unit and weighting an hour's employment of special labour in proportion to its remuneration" to make labor time a homogenous measure. This is similar to Marx's "socially necessary labor time"—see below.
13. In agreement with Hayek, Sraffa wrote of the "well-founded objection to the vagueness of the conception of the general price-level understood as anything different from one out of many possible index-numbers of prices, and in the opinion that such a conception can have no place in a theory of money." (Sraffa 1932, p. 44) However, he chastised Hayek for focusing solely on the medium of exchange function of money (which led Hayek to erroneously embrace the notion of neutral money), arguing that money is "also a store of value, and the standard in terms of which debts, and other legal obligations, habits, opinions, conventions, in short all kinds of relations between men, are

more or less rigidly fixed. As a result when the prices of one or more commodities changes, these relations change in terms of such commodities ...” (Sraffa 1932, p. 43)

14. Forstater (2004, 2005) highlights Marx’s views on the role the state’s taxes played in primitive accumulation and in monetizing labor power and economies.

11. Employment policy and the value of the currency

INTRODUCTION: WHY SHOULD WE GUARANTEE A JOB?

In this penultimate chapter, we examine MMT's advocacy for a Job Guarantee (JG) program. Not only would it eliminate involuntary unemployment, but if correctly designed, it will help stabilize wages and prices. In this chapter, we will argue that stable prices and truly full employment are possible and, indeed, are complements. Further, a JG program will help stabilize the economy by acting in a strong counter-cyclical manner. In this book, we have developed an approach to credit and state money that can stand up to critiques by those who argue that full employment with price and financial stability is impossible.

The United Nations Universal Declaration of Human Rights includes the right to work, not only because work is important in its own right, but also because it underlies many of the other economic and social entitlements that cannot be secured without paying jobs. (Harvey 1999) Joblessness creates a long list of problems—both for individuals and for society as a whole—including self-pity, self-loathing, rage at society (Harvey 2002; see also Harvey 1999); absolute and relative poverty, damage to social status and self-respect, adverse psychological and physical health effects, stress, suicide, crime, and other anti-social behavior. Amartya Sen (1999) supports the right to work because the economic and social costs of unemployment are staggering with far-reaching consequences beyond the single dimension of a loss of income (see also Rawls 1971).

THE PROBLEM IS THAT MARKET ECONOMIES DO NOT PROVIDE JOBS FOR ALL WHO WANT TO WORK.

Economists and policymakers generally fail to acknowledge this, however, and as a result, public policy is often biased against the creation of a sufficient supply of jobs. They argue that full employment is not consistent with price stability. Indeed, unemployment is viewed as a policy tool. As an example, in its post-COVID inflation-fighting stance, the Fed continually argued that the

labor market would have to weaken considerably before it would lower interest rates. In other words, the Fed wanted to see *more* unemployment. If the burden of unemployment in the name of fighting inflation were equally shared, this might be acceptable. However, joblessness is concentrated among groups that suffer other disadvantages: racial and ethnic minorities, immigrants, younger and older individuals, women (especially females heading households with children), people with disabilities, and those with lower educational attainment. Hence, at the very least, safeguards are required to protect those who suffer the consequences of policy that fights inflation by imposing large costs on a minority for the benefits of lower inflation that accrue to society as a whole. (Harvey 2000, 2002; Mitchell and Wray 2005; Wray and Forstater 2004)

The burden of involuntary unemployment and involuntary part-time work is disproportionately borne by those with long spells of unemployment relieved only temporarily by short spells of work—perhaps only part-time work. There are a variety of approaches taken to explain the problem of long-term joblessness, of which the most important views are behavioralist (problems with the individuals who are unemployed), structuralist (for example, skills mismatch), and job shortage. (Harvey 2000) In the US, with the exception of the New Deal experience (when unemployment had reached 25 percent), both public attitudes and policymakers emphasize behavioral and structural problems. This generates policies that try to motivate and train the unemployed, together with the promotion of greater *flexibility* that would reduce labor market *frictions*. However, if the problem is a chronic shortage of full-time jobs, such policies only redistribute unemployment among the unfortunate, who are blamed for their joblessness. Especially in an expansion, those left behind generally do have characteristics identified in the behavioralist and structuralist arguments, hence concealing the true problem—a chronic shortage of full-time employment opportunities.

FULL EMPLOYMENT THROUGH A JOB GUARANTEE

A key justification for government direct provision of jobs is that no capitalist society has ever managed to operate at anything approaching true, full employment on a consistent basis without *direct job creation* on a large scale by the government. (See Mitchell and Muysken 2008; Ginsberg 1983.) Further, the burden of joblessness is borne unequally, concentrated among groups that already face other disadvantages. Finally, *only the government* can offer an infinitely elastic demand for labor (offering to hire all who cannot otherwise find employment). Private firms can only hire the quantity of labor needed to produce the level of output expected to be sold at a profitable price. The

government can take a broader view to include the promotion of the public interest, including the right to work.

Note that this argument is also consistent with the views of Keynes: capitalism as a monetary production economy is structurally unable to achieve and sustain full employment. This is because the purpose of production is to obtain money profits, and the special characteristics of money lead it to act as a barrier to generating sufficient aggregate demand to produce full employment. Keynes viewed this as a chronic problem, not simply a problem in severe recessions and depressions.

While many have interpreted Keynes's argument as supporting a view that government can simply boost demand by increasing its spending—and thereby achieve full employment—Minsky's (1975; 1986; 2013) interpretation of Keynes rejects that as too mechanistic. Simply increasing aggregate demand could move the economy toward full employment, but it will not be sustainable—for three reasons (all of them also of concern to Keynes). First, as the economy approaches full employment, it will tend to generate inflation as the supply elasticity of production across more sectors of the economy falls. Second, Minsky (2013) argued that government spending typically benefits the most advanced sectors of the economy (industry, infrastructure, defense) that are oligopolized and unionized. Both workers and firms take advantage of high demand to push up wages and prices. Not only is this inflationary, but it also increases inequality (profits rise, and higher-paid workers get the wage increases). And finally, Minsky (1978) recognized that government spending policy that favors investment tends to increase financial instability—his *Financial Instability Hypothesis*. Economic growth produces profits and raises expectations, increasing investment and producing even more profits. This can produce a euphoria so that risks are downplayed, and more debt is issued until a financial crisis is created. For these reasons, Minsky rejected the simple “pump-priming” of demand and replaced it with spending targeted to job creation for the unemployed. (See Wray 2015, 2018, 2022a.) He argued that only an Employer of Last Resort (ELR) program could achieve sustainable full employment.

Government should and must play a *direct* role in providing jobs to achieve social justice. Note that the benefits of full employment go far beyond the provision of monetary income. While monetary income can increase consumption by the poor, income alone does not promote social inclusion. While it is true that someone could receive welfare payments and devote time to participating in the life of the community, this is not the same as being paid to do so. Compensation for work performed is seen as quite different from income received through social assistance; and while volunteered work fulfills a desire to contribute to society, the combination of social assistance plus participation

in unpaid work is not equivalent to paid work in the eyes of program participants or of those who work for pay.

The design and justification for MMT's version of the JG follows closely the work of Hyman Minsky in the 1960s and early 1970s when he proposed what he called an ELR program as an alternative to the USA's *War on Poverty* developed by Presidents Kennedy and Johnson. Minsky argued that their package of policies to reduce poverty was missing one essential ingredient: jobs. He argued that without providing jobs with decent pay to all those who want to work, the War on Poverty would not be able to permanently relieve poverty, significantly reduce unemployment, and improve equality.

Minsky modeled his ELR on Roosevelt's New Deal jobs programs, such as the Works Progress Administration and the Civilian Conservation Corps. Those programs helped to relieve suffering during the Great Depression while also transforming America and preparing it for entry into WWII. Workers in the WPA

shouldered the tasks that began to transform the physical face of America. They built roads and schools and bridges and dams. The Cow Palace in San Francisco, La Guardia Airport in New York City and National (now Reagan) Airport in Washington, D.C., the Timberline Lodge in Oregon, the Outer Drive Bridge on Chicago's Lake Shore Drive, the River Walk in San Antonio ... Its workers sewed clothes and stuffed mattresses and repaired toys; served hot lunches to schoolchildren; ministered to the sick; delivered library books to remote hamlets by horseback; rescued flood victims; painted giant murals on the walls of hospitals, high schools, courthouses, and city halls; performed plays and played music before eager audiences; and wrote guides to the forty-eight states that even today remain models for what such books should be. And when the clouds of an oncoming world war loomed over the United States, it was the WPA's workers who modernized the army and air bases and trained in vast numbers to supply the nation's military needs. (Taylor 2008, p. 2)

The New Deal jobs programs employed 13 million people, with the WPA being the biggest program, employing 8.5 million, lasting 8 years and spending about \$10.5 billion. (Taylor 2008, p. 3) The WPA built 650,000 miles of roads, 78,000 bridges, 125,000 civilian and military buildings, and 700 miles of airport runways; it fed 900 million hot lunches to kids, operated 1500 nursery schools, gave concerts before audiences of 150 million, and created 475,000 works of art. (Taylor 2008, pp. 523–4) While of course the USA already had some big cities and important industries, most of the country was underdeveloped when Roosevelt assumed office. Americans today do not fully recognize the role played by those jobs programs in transforming America into a modern developed—and rich—country.

There have been many job creation programs implemented around the world, some of which were narrowly targeted while others were broad-based.

In the aftermath of its economic crisis and the collapse of its currency board, Argentina created *Plan Jefes y Jefas*, which guaranteed a job for poor heads of households. (Tcherneva and Wray 2007; Tcherneva 2020) The program successfully created 2 million new jobs that not only provided employment and income for poor families, but also provided needed services and free goods to their neighborhoods. The program was highly decentralized with an innovative approach to project approval and administration. It can serve as an example of good practices (as well as for learning from mistakes).

In 2005, India passed the National Rural Employment Guarantee Act that commits the government to providing employment in a public works project to any adult living in a rural area. The job must be provided within 15 days of registration and must provide employment for a minimum of 100 days per year. (Hirway 2006) It quickly became by far the biggest job guarantee program in the world (over 50 million workers in 2010), although it is not a universal program, is only part-year, and implementation has been uneven as local authorities are not always supportive. However, it is unusual because it provides both a right to a job (for those who qualify) and also mandates the right to information, so that authorities must tell the public about their right to demand a job.

More recently, the United Nations' Human Rights Council requested its Special Rapporteur "to examine and report back to member States on initiatives taken to promote and protect the rights of those living in extreme poverty, with a view to advancing the eradication of such poverty." (<https://www.ohchr.org/en/special-procedures/sr-poverty>) As part of the investigation, the Special Rapporteur sought input on the role that a JG can play as "a tool in the fight against poverty." One of its reports emphasized that growth alone would not resolve the problem of poverty:

The dominant approach to the fight against poverty relies on increasing the aggregate output of the economy (measured as the gross domestic product), combined with post-market redistribution through taxes and transfers. The Special Rapporteur argues, however, that the current focus on increasing the gross domestic product is misguided. An increase in gross domestic product is not a precondition for the realization of human rights or for combating poverty and inequalities. The ideology of "growthism" should not become a distraction from the urgent need both to provide more of the goods and services that enhance well-being and to reduce the production of what is unnecessary or even toxic. As long as the economy is driven mainly by profit maximization, it will respond to the demand expressed by the richest groups of society, leading to extractive forms of production that worsen social exclusion in the name of creating more wealth, and it will fail to fulfill the rights of those in poverty. Moving from an economy driven by the search for maximizing profits to a human rights economy is possible and, to remain within planetary boundaries, necessary.¹

The report emphasizes that growth can exacerbate inequality and environmental damage without reducing poverty. An earlier report recommended the JG as an alternative to “growthism”:

The Special Rapporteur on extreme poverty and human rights explains how the introduction of a job guarantee, with the State acting as an employer of last resort, can contribute to the full realization of the right to work, transforming it from a policy objective into an enforceable legal right. The job guarantee tackles the paradox of structural unemployment and underemployment coexisting with important unsatisfied societal needs, as neither the State nor the market currently supplies the public goods that are needed for the greening of the economy and for a thriving care economy. The Special Rapporteur highlights the benefits of the introduction of the job guarantee both to the individual and to the community, and addresses a number of objections to the idea. He sees the introduction of a job guarantee as an essential component of the “just transition” and of the new eco-social contract needed for the post-crisis recovery.²

In the next section, we look at a JG proposal for the USA.

A JOB GUARANTEE PROGRAM FOR THE USA

Since WWII, it has been the official policy of the US government to simultaneously pursue high employment and stable prices. These goals are the subject of two laws, the 1946 Employment Act and the 1978 Humphrey–Hawkins Act; the latter Act strengthened the government’s commitment to employment by setting a goal of full employment, with an interim goal of an adult unemployment rate of 3 percent (or 4 percent overall). Paradoxically, many, perhaps most, economists doubt that it is even possible to achieve anything close to a 3 percent unemployment rate without at the same time inducing accelerating inflation. As a result, for at least the past five decades, monetary policy generally has been geared toward raising the unemployment rate to achieve stable prices—most recently demonstrated by Chairman Powell’s response to the COVID-induced inflation. Unemployment is widely perceived as the inevitable cost of price stability.³

We reject that and argue that the government can guarantee a near-zero measured unemployment rate, meaning that all who are ready and willing to work at the going wage will be able to find a job—only those unwilling (or unable—a point we will address below, as we believe that all people with disabilities should be accommodated if they want jobs) to work at the going wage would be left without work (and these are not normally counted as unemployed). At the same time, by setting this going wage, the government will provide a price anchor to impart greater price stability to the system.

We do not claim that this policy would cause any particular price index to remain constant over time (recall the discussion in the previous chapter of

the necessarily arbitrary choice of an index). The proposed policy would still allow market (and other) forces to affect both nominal and relative prices—which might be essential to allow markets to “work” (to equalize rates of profit and exploitation, as discussed in Chapter 10). However, the point is that the proposed *full employment policy* would not generate the sort of inflationary pressures that many economists believe must result from high employment. Thus, inflation—defined as persistent increase in some price index—could coexist with our proposed policy but would not be caused by the policy. That is, we will show that a true full employment policy is not, in itself, *inflationary* and indeed might reduce inflationary pressures. Further, the policy would help to reduce economic fluctuations (the *business cycle*) through a powerful built-in automatic stabilizer (although we make no claim that this would be sufficient to eliminate business fluctuations, nor are we certain that such would be desirable).

In the next two sections, we turn to the two primary components of the proposal. The government would: (a) act as employer of last resort, and (b) set the base wage throughout the economy. In later sections, we will examine the general theoretical background and other implications of the proposal.

Before presenting the proposal, however, we will first list some caveats:

1. Those who are not ready and willing to work are not the intended targets of this program. It is not a welfare program.
2. No one will be forced by the government to work in JG. The program is only for those who are willing to participate. JG is not a form of workfare—that is, it will not force recipients of welfare to work for their benefits. JG can supplement any welfare safety net that is politically acceptable.
3. JG workers can be fired, with restrictions placed on rehiring. Those who do not measure up to required standards will be fired; they should be given a second or third chance, but at some point, they will have to rely on the social safety net.
4. The JG wage will be a substantial improvement over the average package of welfare and unemployment benefits received in the US; in other countries with more generous safety nets, the improvement may be less significant. The following discussion should make it clear that the essentials of our analysis hold regardless of the initial setting of the JG wage.
5. We recognize that JG alone cannot resolve all employment, unemployment, underemployment, low income, and disability problems. We do believe that it offers a major improvement over the current situation. However, other social programs as well as labor force programs (apprenticeships, technical training, and so on) will be required even after JG is put into place.

In the next few sections, we will examine the general set-up and theoretical impacts of such a program before summarizing the results from a simulation of the effects on the US economy of implementing a JG program. (For more details on the administration and kinds of jobs that would be included in a JG program, see Tcherneva 2020.)

THE JOB GUARANTEE PROPOSAL: STRUCTURE

The first component of the proposal is relatively simple: the government acts as the *employer of last resort*, hiring all the labor that cannot find private or public sector employment. As Hyman Minsky said:

The policy problem is to develop a strategy for full employment that does not lead to instability, inflation, and unemployment. The main instrument of such a policy is the creation of an infinitely elastic demand for labor at a floor or minimum wage that does not depend upon long-run and short-run profit expectations of business. Since only government can divorce the offering of employment from the profitability of hiring workers, the infinitely elastic demand for labor must be created by government. (Minsky 1986, p. 308)

We have called it *Public Service Employment* (PSE) in the past, as “last resort” has negative connotations and Bill Mitchell called it *Buffer Stock Employment* (for reasons to be discussed), but most advocates have settled on *Job Guarantee* (JG), which is the term we will use. The government announces the wage and hires all who seek employment at that wage. Of course, there will remain many (non-JG) jobs in the public sector that will generally pay wages above the JG wage. It is also important to emphasize that the JG is not meant to substitute for current public sector employment (JG workers should not displace current public employees).

For the sake of our discussion, we assumed that the program’s wage would be \$15 per hour or \$30,000 per year for full-time (JG) employment. This is about double the Federal minimum wage in 2024, although many states and cities have higher minimum wages—with many approaching our proposed wage. This is the wage that we used for economic simulations for a 2018 Levy Economics Institute study (the results of which will be summarized below), so it is convenient to use it here. The program wage should be a *living wage* appropriate for the living standards of the nation. Although it may be necessary to begin with a lower wage (depending on the wage structure of the country), eventually, it should be raised to a level sufficient to provide a decent living standard—which is consistent with the goals of the UN Charter (cited above). As a nation’s productive capacity rises, the program wage should also rise so that its workers obtain a share of the increased output due to rising productivity.

Since the program pays \$15 per hour, that will become the effective minimum wage all over the country, as workers will not take private or public sector jobs that pay less unless there are other perceived benefits to doing so. (As Hyman Minsky always argued, if there is any unemployment, the *effective* minimum wage is zero; minimum wage laws are *effective* only at full employment. A JG program enforces an effective minimum wage even without a minimum wage law.) Note that some workers might prefer part-time jobs, and there is no reason not to allow this, but to simplify discussion and analysis, we will assume here that JG employment is full-time.

This policy will eliminate involuntary unemployment and underemployment, defined as workers willing to work at the going wage but unable to find a full-time job. Many individuals—even those in the labor force—may choose not to accept the job offer: there will be those who are unwilling to work in the program (perhaps at any wage), those who are unwilling to work for the program's wage, those who are between jobs and prefer to look for a better job while unemployed, and so on. It is not optimal (either socially or individually) for everyone to be fully employed—voluntary unemployment (especially short-term) can be rational. Thus, our concern is to ensure that all those ready, willing, and able to work at the program wage will be able to obtain a job. We define this as a state of full employment.

One implication of the JG is that much social spending currently targeted to low-income people might be reduced, such as federal and state income support, food stamps, housing assistance, and unemployment compensation—simply because individuals do not need them or no longer qualify because their income is too high. Note that we are not advocating the elimination of these social programs, but spending on them will decline.

Setting the JG wage above the current minimum wage—and above private and public sector wages currently paid by many employers—would force them to respond with wage and/or benefit increases to retain their workforce. Would the higher wages cause inflation? Furthermore, adding millions of jobs will likely increase aggregate demand. Could this cause inflation? The economic implications of creating a JG program and raising wages are the subject of the next sections.

THE JG WAGE AND EFFECTS ON WAGES AND PRICES

One fear is that adding millions of jobs—at wages higher than the current minimum wage—through a JG program will increase consumption and cause inflation. Here we address that concern.

The number of JG workers will be determined by the performance of the economy, with participation moving in a highly countercyclical manner. When the economy slows, private firms (and some public employers) lay off workers,

who move into the JG program; when the economy recovers, they will be rehired from the JG program. Government spending will likewise move countercyclically since spending on the JG program (as well as on other social assistance) goes up in recession and down in expansion. (Recall the discussion of the Wray Curve in Chapter 8: adding a JG program would make social spending more countercyclical.) On the other hand, tax revenue moves procyclically (strongly so in the USA)—taking more income out of the economy in expansion and much less in recession. With a JG, the government's budget deficit will be strongly countercyclical. This is precisely what is desired if the goal is to use the government's budget to help stabilize the economy—the functional finance view.

The swings of employment in the JG program as well as the movement of the government's budget stance will be largely determined by the private sector—not by the government. When households and firms reduce spending, government spending rises to accommodate their desire to save more. We see this effect in the sectoral balances: private sector balances move sharply to greater surplus in recessions and toward smaller surpluses (and, more rarely, even into deficits) in expansions—and the government's balance moves to offset that. (In a closed economy, we would observe a perfect mirror image because there would be no foreign sector; in the case of the USA, there is always a deficit in the current account, so the government deficit is matched by the sum of the private surplus and foreign surplus.) The implication is that adding a JG program will reduce swings of the business cycle by increasing the response of the government's balance to swings of private sector spending. Aggregate demand will follow a smoother path through time—less inflation pressure in expansion and less disinflationary pressure in recession years. This helps to stabilize prices.

Before WWII, the USA experienced both inflation and deflation—there was no general trend of prices over long periods of time, although prices could rise sharply (usually during wars) and fall sharply in depressions. The post-WWII period is quite different—while inflation comes and goes, outright deflation has largely been banished in the rich countries (with Japan a notable exception). Adding the JG program would probably make it even more unlikely to experience deflation, as involuntary unemployment would be eliminated. While that means the general trend of prices will remain upward, it eliminates the devastating effects of deflation and debt deflation—which is what caused the Great Depression to be so severe and long-lasting.

In sum, the JG program will help to stabilize aggregate demand and—especially—to cushion downturns that in the distant past used to lead to depressions and deflation. For reasons we discuss next, the JG program will probably reduce inflationary pressures in robust expansions—reducing peak inflation rates. But the net impact might be that the price level continues to rise over

time—an inflation “ratchet”: moderate inflation in good times, no deflation in bad times. However, this is compensated by more economic stability—a trade-off that the rich countries seem to have accepted throughout the postwar period as price levels generally rose, but the downturns were less economically destructive. Adding the JG eliminates involuntary unemployment in both good times and bad times.

There is one more issue to cover before turning to wage pressures. When we first implement the JG, we could be creating millions of new jobs at \$15 per hour. That would raise consumption and aggregate demand, possibly creating price pressure. This is a “one-off” impact.⁴ The economy would adjust as the higher demand would encourage investment in productive capacity. However, note that this increase in aggregate demand can be avoided, if desired, by simultaneously imposing new taxes or reducing government spending elsewhere. We are not recommending this, but the point is that a generous JG program can be implemented while even reducing aggregate demand through a compensating fiscal adjustment. The increased consumption by JG workers will be offset by the reduced consumption of other workers (paying the new taxes) as well as by those who receive profit or rental income—this is essentially a redistribution of output toward those who were previously involuntarily unemployed. That could even be a desired outcome.

The point is that adding a JG program does not necessarily mean that aggregate demand will be increased. Whether a tax hike (or reduction of other spending) should be adopted depends on the assessment of productive capacity and possible (transitory) effects on prices of moving to full employment through a JG program. As should be obvious from the discussion in earlier chapters, a tax hike is not needed to “pay for” the JG program but might be desired if productive capacity is too limited. It is a pragmatic decision. Below we will see that our simulation of a JG program indicates that there is little reason to depress aggregate demand as the inflation pressure is limited and transitory—at least in the case of the USA and most other developed nations that normally operate with excess capacity.

Note, however, that with a JG in place, austere fiscal (or monetary) policy cannot cause unemployment—it simply moves more workers into JG employment and out of the private sector. The dynamic response of the economy to policy changes with a JG: tight fiscal and monetary policy shifts employment to the public sector and reduces the growth of the private sector but does not increase unemployment.

The initial impact of the JG on wages depends on where the wage is set. If the JG wage is set at or just above the existing statutory minimum wage, and if this minimum wage had universal coverage before JG, then low-wage private sector jobs will experience only minimal impacts—private wages need rise only sufficiently to make private sector employment preferable to the JG.

So, at the low end of the wage scale, implementation of JG might cause wages and prices of products produced by these workers to experience a small one-time increase. However, this one-time jump—no matter how large it is—is not inflation nor can it be accelerating inflation as these terms are normally defined by economists (inflation is defined as a continuously rising price level). Furthermore, a policy that raises wages at the bottom should be desirable—and is similar in impact to an increase in the legal minimum wage.

However, we recommend setting the initial wage at \$15 per hour—double the Federal minimum in 2024. What about wage pressures in that case? The implementation of the JG would raise wages outside the program—just as an increase of the minimum wage would (by design) raise wages at the bottom. What is the implication for other employers and for wage-led inflation?

With the program wage at \$15 per hour, other employers who pay lower wages (and/or who provide lower benefits) will need to boost theirs or will lose their workers. Some employers may not be able to pay higher wages because they cannot pass higher labor costs on to customers. Their workers can move into the JG program. Some kinds of products or services might disappear from the market as producers cannot cover the higher wage costs, or they might be provided through imports from low-wage countries. Some employers might be able to replace workers with “robots”; others might be able to innovate to find less labor-intensive production processes. Some low-wage firms will go out of business.

All of these are also possible consequences of raising the minimum wage, of course. It is often claimed that raising minimum wages causes unemployment, although the evidence runs strongly against this belief. (Card and Krueger 1994) It seems that even if some firms must close shop, the demand effects produced by higher wages produce enough new employment to offset those losses. In the case of the JG, employment is assured even if some firms shut down. If raising minimum wages—and the JG wage—causes the exit of firms, this can be justified on the argument that business models that require wages lower than an acceptable minimum standard should not be permitted to persist. Alternatively, if deemed to be in the public interest, the goods or services supplied by them can be publicly supplied or subsidized.

The initial jump in wages due to the implementation of a JG program is a one-off adjustment as long as the JG wage remains at \$15. Employers will then be forced to cover these higher costs through a combination of higher product prices, greater labor productivity, and lower realized profits. Thus, some product prices could also experience a one-time jump as the JG program is implemented. At some point, the JG wage will need to be raised because, as discussed, price levels trend upward (with or without a JG program). We recommend periodic adjustment of the JG wage but do not favor automatic indexing.

As we discuss below, our simulation finds that inflation pressure quickly dissipates even if the JG is implemented at \$15 per hour. Let's turn to the claim that the existence of the JG will embolden workers to demand higher wages when the threat of unemployment is removed.

WAGES AND A RESERVE ARMY OF THE EMPLOYED

Some critics have argued that wages are likely to continue to rise because, by achieving full employment of labor, the threat of unemployment is removed, emboldening workers to demand continual wage hikes, leading to a wage-price spiral. This is essentially the old Marxist *reserve army of the unemployed* argument. Workers who might have previously earned \$30,000 per year now demand \$35,000, knowing that, in the worst case, they might be fired if they are too obstinate—but they can go into the JG program and earn \$30,000. By extrapolation, all workers might harden their positions, causing wages to jump upward. Prices would move upward to the extent that higher labor productivity and lower profits could not absorb the entire increase in wages.

However, again, this is a one-time jump that is not defined as inflation, unless it can be argued that all workers above the \$30,000 threshold continuously raise wage demands over time (generating a *wage-price spiral*). This appears unlikely. The marginal \$30,000-a-year worker who decides to demand \$35,000 per year, calculating that this is worth the risk of losing their job (to take JG jobs), will not face the same decision once they are earning \$35,000 and demanding \$40,000—for now, the loss is potentially \$10,000 per year. It is hard to see how the guaranteed job will cause any individual worker to continually increase the wage demand over time, because the further from the JG wage benchmark, the greater the loss due to obstinacy.

In orthodoxy, the individual's wage is supposed to equal the *marginal productivity* of the worker—those with higher productivity receive higher wages. Both of these are perceived to be in “real” terms—the “real” productivity of the marginal worker is equated to the “real” wage. However, as discussed in the last chapter, there are only two measuring units that can be applied to heterogeneous production: labor hours and money value. The goal of production is not the production of some physical commodity but money value—profits measured in money terms. Profit comes out of surplus labor, determined by the degree of exploitation. There is a tendency to equalize the rate of exploitation. In place of marginal productivity, we should substitute the ability of the employer to squeeze monetary profits out of workers (what Marx called exploitation). The higher the money wage, the less profitable the employment of workers becomes—given some maximum to the *speed-up* (making workers work harder and faster), and a limit to the firm's ability to raise prices.

In what follows, by *productivity* we mean the generation of gross money profits—value of sales less wages paid to produce the revenue. Note that we are not arguing that productivity is to be found in the nature of the worker, but rather (largely) in the nature of the employment. (Of course, skill level, experience, education—as well as discrimination by race and gender—all play a role in determining who gets what job.)

Money revenue per worker varies greatly across employments, and wages will tend to be lowest in sectors of the economy where revenue generated per worker is lower. However, the *rate of exploitation* will tend toward equality. For example, assume that the rate of exploitation is 50 percent, so on average, workers in each sector produce revenue in four hours that is equal to the wage paid for eight hours of work. A job that pays a wage of \$5 per hour generates \$80 of revenue per day per worker, of which \$40 covers the wage and the other \$40 goes to gross profits (gross profits must cover interest, taxes, rent, and other business costs). Meanwhile, a job that pays \$20 per hour generates \$160 to cover wages and \$160 for gross profits. We can say by definition that the second worker is more productive than the first.

The government's JG wage becomes the base wage in the economy—the wage for the pool of lowest productivity workers in the private sector during periods of normal demand. Those workers whose productivity is substantially higher will be in private sector jobs at higher wages. (These are tendencies—rates of exploitation tend to be equalized across employment. Note also that discrimination—by race, gender, etc.—plays a role in the deviation of rates of exploitation.) When private demand is below normal, workers who formerly were more productive and earned higher wages move into the JG employment. When private demand is higher, JG workers leave the pool and move into higher productivity work, earning higher wages.

At high enough demand, workers can move up the employment ladder to ever-higher productivity jobs with higher wages, and wages generally might rise as employers try to retain workers. However, just as workers have the alternative of JG, employers have the opportunity of hiring from the JG pool. This is the primary *price stabilization* feature of the JG program. If the wage demands of workers in the private sector exceed by too great a margin the employer's calculation of their productivity, the alternative is to obtain JG workers at a mark-up over the JG wage. The JG pool will operate as a *buffer stock*—as discussed below—and just as a buffer stock of any commodity can be used to stabilize its price, the government's labor buffer stock will help to stabilize the price of non-JG labor to the extent that workers in the JG pool are substitutes for non-JG labor.⁵

The JG workers are not lost as a reserve army of potential employees; rather, they can always be obtained at a mark-up over the JG wage. In the absence of JG, these workers can be obtained at a mark-up over the value of the package

of social spending and private income obtained when unemployed (unemployment compensation, food stamps, under-the-table work, handouts, and so on); this markup, however, is likely to be higher than the markup over the JG wage since it must be sufficient to make employment preferable to unemployment (and recall that those in the JG pool have demonstrated that they are ready, willing, and able to work).

Further, perceived employability falls as the length of spell of unemployment rises (long-term unemployment greatly reduces the probability of obtaining a job); with the JG available as an option, involuntary long-term unemployment will fall. JG workers demonstrate that they are ready and willing to work in a manner that the unemployed cannot. Furthermore, policy can be geared toward enhancing the skills of the JG pool—Minsky argued that every job in the JG program should include training. This would make JG workers more desirable and increase the rate at which they are hired into the private sector, thereby diminishing inflationary pressures in periods of high demand.

It is hard to imagine that true full employment with a JG program would be more inflationary than the current system that relies on unemployed labor and excess capacity to dampen wage and price increases; however, it pays unemployed labor for not working, and in some cases, those who are unemployed develop behaviors that act as barriers to private sector employment. Social spending on the unemployed prevents aggregate demand from falling excessively, but little is done to promote aggregate supply and growth of potential output. With JG in place, however, labor is paid for working, which can lead to the production of publicly supplied goods and services, can promote efficiency of the private sector (if, for example, JG generates productivity-enhancing public and social infrastructure) and reduce private sector costs (for example, by reducing crime), and can increase the education and skills of JG workers (compared with education and skill levels of the unemployed). Thus, JG might increase aggregate supply (or potential output) and thereby place downward pressure on prices, rather than causing inflation.

This appears to be an unusual claim: full employment might be disinflationary? We expect that those who engage in JG work will become more employable, and thus operate as a better *reserve army of the employed* than the unemployed. Further, some of the private and social spending associated with long-term unemployment, including homelessness, health problems, and crime, will be reduced. It is conceivable that by reducing spending associated with the social costs of unemployment, achieving full employment through a JG could lower overall aggregate demand (even without the fiscal austerity discussed above). This can be mitigated by reducing taxes and/or increasing non-JG government spending if necessary.

There are two conclusions that follow. If JG is put in place, it is unlikely that this will be inflationary in the sense of generating continuous pressure

on wages and prices. Wages would experience a one-time increase—to at least \$15 per hour—and workers of higher productivity might become more obstinate in their wage demands, causing other wages also ratchet upward. However, against this tendency, the JG will replace unemployment, increasing the supply of labor deemed employable. When demand for private output rises sufficiently that additional workers are needed, JG can supply them. Thus, even the one-time jump in wages and prices might be quite small.

THE JG AS A BUFFER STOCK PROGRAM

Some economists (including Keynes) have noted that there really is no labor market; labor is not like other commodities because it cannot be owned, it cannot be stored, and it is not nearly as mobile as other factors of production typically are. (See Galbraith, 1997; and a humorous piece by Prasch (2003), who explained why labor is not like broccoli.) In addition, there are considerable *information costs* and uncertainty involved in hiring workers. The JG program will resolve or reduce some of these difficulties. In a sense, the JG allows *storage* of labor—when it is not needed by private employers (and non-JG public employers), it can be stored in the JG buffer stock pool. JG employees will be able to provide employment records to potential employers; if JG is well administered, workers will also have records of skills and on-the-job training obtained in JG.

Both Hyman Minsky (2013) and Bill Mitchell (1998) have compared the JG to a commodity buffer stock program—like the agricultural price support programs in the USA (corn) and Australia (wool), in which the government maintains ceiling and floor prices. With a JG, the government acts as a market-maker, standing ready to *buy* unemployed labor at a fixed price, or to *sell* (provide it to non-JG employers) at a mark-up over the JG wage. As in all buffer stock schemes, the commodity used as a buffer stock is always fully employed.⁶ It also always has a stable price, which cannot deviate much from the range established by the government's announced buy and sell prices. What we are proposing to do is to make a market in labor by establishing a *buffer stock of labor*. This permits full employment and stable prices.

From time to time, there will be pressure for an upward revision of the JG wage. As the overall price level will not be held constant, and as substantial forces in modern capitalist economies generate trend increases of the price level, the *real* (inflation-adjusted) JG wage will tend to fall over time—generating a need for an adjustment. In addition, there will be obvious pressures by labor to raise the JG wage—just as there are pressures currently to increase the minimum wage. When the government raises the JG wage, this in effect devalues the currency by redefining what must be done to obtain the currency as it buys fewer labor hours.

In the previous chapter, we have followed Foley in arguing that the value of money is determined by the monetary equivalent of labor effort by “simple, abstract, socially-necessary labor” or what Keynes called “ordinary labor.” For example, the decision to compensate for inflation by increasing the hourly wage from \$15 per hour to \$20 per hour reflects a 33 percent devaluation of the value of the currency. Again, other wages (and prices) will also adjust upward to reflect the devaluation—but there is no reason to suppose that this will be *inflationary* in the sense of a continuing rise in prices. Rather than causing inflation, the devaluation takes account of inflation that results from factors that have little to do with the JG policy.

This would be similar to a devaluation of the currency under a gold standard, in which the government acts as a market-maker for gold, utilizing a gold buffer stock. The government sets the price of gold in the domestic currency, then makes a market in gold. The government’s price (for example, \$32 per ounce of gold) is fixed, while its spending on gold is *market-determined* (as it buys all gold offered for sale at that price). It has long been argued that a gold standard imparts some stability to prices; however, even on a gold standard, prices may rise and might induce the government to devalue (raise the money price of gold).⁷ The analogy with JG should be clear—devaluing the currency by raising the price of gold or the wage of JG workers can simply be a response to inflation arising from other sources.

Note also that just as a gold standard ensures that gold is always fully employed (gold can be sold to the government at the fixed price if it is not desired as a hoard), the JG *labor* standard ensures full employment of labor. Under JG, the government would *monetize* labor just as it monetizes gold under the gold standard. (Note, we reject the notion that gold gives money value—rather, on a gold standard, the government chooses to give gold a fixed monetary value in terms of the currency.)

The question is: Do we prefer to have gold fully employed, or is it preferable to have labor fully employed? Should we attempt to stabilize prices by re-establishing the gold standard, or should we move to a JG wage standard? Or should we simply continue the present path, which requires the unemployment of labor to try to minimize inflation?

To the extent that the labor buffer stock helps to reduce wage pressure, it helps to reduce inflation. As Kahn put it, in his *General Theory*, Keynes had adopted what Hicks called the wage theorem:

The money-wage is the fulcrum on which rests the whole structure of everything expressed in terms of money—all prices, incomes of every kind, and all money-values. A higher level of money-wages means that everything expressed in terms of money is higher in the same proportion. (Kahn 1974, p. 16)

Kahn goes on, stating that if the money wage rises,

[A]ll incomes and debts fixed contractually in terms of money are smaller in real value as a result of the money-wage being higher ... The basis of the fundamental role of the money-wage in determining all prices, money-incomes, and money values is that money-wages not only form part of costs of production but, because they are to a large extent spent, they form part of total purchasing power expressed in terms of money. (Kahn 1974, p. 17)

It makes more sense to operate with a labor standard than with a gold standard—as labor goes into the production of all goods and services, while gold is an input to few production processes. Fixing the money value of gold will not stabilize the price of produced output because it is not an important input to most production processes. Wages *are* a major input to producing almost everything.

The buffer stock aspects of JG generate *loose* labor markets (in the sense that there is an increase in the number of workers deemed to be employable) even as they ensure full employment. This stands in stark contrast with “Keynesian” demand management policies that were designed to *prime the pump* with aggregate government spending that would increase private demand sufficiently to lower unemployment to the full employment level. Minsky argued that this would lead to *tight* labor markets due to bottlenecks in rapidly advancing or high-productivity sectors that would pull up the entire wage structure so that inflation would be generated long before full employment could be reached.⁸ Indeed, most economists today believe that Keynesian policy proved to be a failure precisely because the tight labor markets *did* generate unacceptable levels of inflation.

The JG is not subject to the same criticism, for it allows *loose labor markets* with a pool of labor that can be hired, even at full employment. If the JG pool shrinks too much during an expansion so that it cannot act as a buffer stock, the government can either raise taxes or reduce non-JG spending to replenish the buffer stock. This allows the private economy to grow at its non-inflationary rate without requiring unemployment since any labor shed by the private sector is absorbed into the JG pool. Thus, aggregate fine-tuning would operate through increases or decreases of the buffer stock, rather than by causing unemployment. While we doubt this would be needed, it is conceivable that a runaway expansion might reduce the JG pool excessively, requiring a fiscal response (reducing other spending or raising taxes). This would still be a possible policy response—just as it is today—although it should be less likely to be required given the JG’s automatic counter-cyclical nature.

As we have argued, the JG will achieve what most economists would call zero unemployment (well beyond what they would call full employment) without creating inflationary pressures. The government will define the currency

by setting the price of standard labor. This might initially devalue the currency—and periodic redefinition by raising wages to compensate for rising prices is likely. In conclusion, the JG policy is not likely to induce inflation—much less to cause accelerating inflation—even if it does cause prices to rise when implemented and each time the JG wage is raised. However, the price anchor of the JG wage will impart a greater degree of stability to wages by setting a well-known wage for homogeneous, standard buffer stock labor that can be used by private employers as an alternative to higher-skilled workers with market-determined wages. And if it became necessary, a fiscal response to dampen demand would increase the size of the JG buffer stock—without causing unemployment. That might make the fiscal response more politically palatable than it is currently.

Modern nations have abandoned the gold standard and adopted a fiat money standard, in which the liabilities of the government (in the US, Treasury coin, Fed notes and bank reserves) serve as the *ultimate money*. Some countries peg to the dollar, making it the ultimate money; within the Euro area, individual nations dropped their own currencies, making the euro the ultimate money. The government spends by issuing currency (mostly in the form of central bank reserves). If a government can create at will the money that the public offers goods and services (especially labor services, for our purposes here) to obtain, then the government's spending is never constrained by narrow financing decisions. The government can offer to hire all unemployed workers at any wage it chooses, allowing government spending (and the government deficit) to float as high as necessary to ensure that unemployment is eliminated.

Still, a JG will not eliminate the business cycle. When private expectations are low, desired net nominal saving rises, and some \$40,000 per year workers lose jobs, moving into the JG to earn \$30,000 per year. The government's spending increases but not by as much as the private income that is lost. The combination of a rising deficit (thus supporting private saving) and falling aggregate income restores an equilibrium level of effective demand (where the private sector hires the number of workers consistent with producing output that is expected to be sold profitably) without unemployment—but at a lower level of aggregate activity. When private demand expands (desired private saving falls), leading to the creation of higher-wage private sector jobs, these replace JG jobs, raising aggregate income. Effective demand is re-established at zero unemployment, a lower deficit, lower actual and desired private savings, and a higher level of output. Thus, the business cycle persists, but with a smaller amplitude.

THE LEVY INSTITUTE'S PROPOSAL FOR A UNIVERSAL JOB GUARANTEE PROGRAM

In 2018, Levy Economics Institute scholars laid out a proposal for a universal JG program for the USA. We then used the well-known Fair Model to simulate the economic effects of the program. First, we will summarize the principles and assumptions behind the proposal, and then we will examine the economic effects. Interested readers should look at the full report for more details. (See Wray 2018, and Wray, et.al., 2018.)

Principles and Assumptions Used

1. The JG should pay a living wage with good benefits. In line with other progressive proposals made at the time, the JG wage would be set at \$15 per hour, with free Medicare-style healthcare. It should also provide free childcare to enable parents to participate in the program. The JG program itself can greatly expand childcare provisioning, as many JG workers can be employed to provide childcare to underserved communities and to JG workers.
2. Congress will appropriate the necessary funds to pay program expenses. No additional taxes will be levied. Advocates of the JG should not play the “how are you going to pay for it?” game. Sovereign currency nations can make all payments as they come due. Taxes should be raised only if the program increases aggregate demand beyond capacity. However, our simulation shows that inflationary pressures will be minimal.
3. The JG should be universal—providing paid work to anyone ready and willing to work. It should also be universal in the sense that it serves every community, offering jobs where people live and providing real benefits to their communities. The JG's wage and benefits will set the minimum standard nationwide, boosting communities across the nation, with relatively greater benefits where they are needed most: where jobs are scarce, where pay is lowest, and where markets suffer the most from lack of income to support buying power. Given the projected size of the program, it will probably be necessary to phase it in over several years. Our simulation assumed a quick, one-year phase-in.
4. The JG should not devolve to either workfare or welfare. The social safety net should not be dismantled; no existing social services should be eliminated. Individuals should be able to continue to receive existing benefits if they do not want to work in the JG program, and any benefits for which they still qualify when they join the program. At the same time, the JG should not provide income support to those that do not work in

the program. The program should have visible benefits to communities so that the workers in the program are recognized as making positive contributions in return for their wages. Workers can be fired for cause—with grievance procedures established to protect their rights, and with conditions on rehiring into the program.

5. However, there should be room in the JG for time-limited training and education. While on-the-job training should be a part of every project, proposals can be solicited for specific training and basic education programs that will prepare workers for jobs in the JG—and, eventually, for work outside the JG. It is important that these are time-limited and that the training is for jobs that exist.
6. Project implementation and management will be decentralized. While funding comes from the Federal government, there should be diversity in the types of employments and employers—to help ensure there are projects that appeal to workers and their communities. Projects should go through several layers of approval before implementation (local, state or regional, federal) and be evaluated at these levels once in operation. Decentralization helps to protect the program from whatever political winds emanate from the *du jour* occupants of the White House and Congress.
7. Where possible, proposals should scale-up existing projects with proven track records and with adequate administrative capacity to add JG workers. Making use of existing capacity will minimize additional overhead. Program funds should be focused on paying wages and benefits to JG employees. Federal spending should not subsidize administrative expenses but can be used for limited purchases of supplies. Typically, JG-type programs devote 25 percent of spending toward tools and other materials.
8. The JG should not be used to subsidize wages of workers employed by for-profit firms. This distorts markets and is not likely to generate substantial new employment. However, private firms will benefit indirectly (and greatly) from the program as it provides a pool of hireable labor and as it contributes to economic growth that improves markets for firms.
9. Direct employment by the Federal government should not dominate the program. Most employment should be administered at the local level—where the workers are, in the communities where they will work. The JG program will probably need to create 15 million new jobs—six times greater than the number of Federal employees. If all 15 million were to join the Federal workforce, supervision of these new workers would, alone, require hiring a large number of additional Federal employees. The Federal government's role in the direct provision of jobs should be focused on providing projects to underserved communities and workers—after

not-for-profits and state and local governments have employed as many as they can.

10. Inclusivity and experimentation should be encouraged. The Federal government should solicit proposals for novel approaches to job creation. For example, workers' co-ops could be formed to propose projects in which wages, benefits, and limited materials costs would be covered by the Federal government for a specified time, after which the co-ops would be self-sustaining, folded into the JG program, or abandoned.
11. Consistent with point 10, project proposals should not be summarily dismissed due to political bias. The JG program should welcome diversity. We should entertain the notion that even those at opposite ends of the political spectrum might have good ideas for projects. Determination of an organization's eligibility to submit projects would be similar to the process the IRS uses for designation of tax-exempt status, although the standards for keeping politics and religion out of the projects should be higher, since the Federal government will spend directly for employment in the program.
12. With decentralization, the types of projects permitted would take account of local laws and rules, including prevailing wage laws and union wage rates. With the JG paying \$15 per hour, in many states and localities, rules and laws will prohibit various types of work, including construction (because these should use union labor or at least pay union wages). In those areas, JG workers will not build infrastructure, for example. It is possible that limited-term training or apprentice projects could be funded in those areas instead. However, in many states construction by JG workers paid \$15 per hour would be permissible. Decentralization helps to ensure conformity with local laws and rules, while maintaining a uniform JG wage.
13. Exceptions to the uniform wage should be considered but should not become the norm. For example, state or local governments might want to subsidize (at their own costs) the Federally paid wage of \$15 per hour to increase wages to some higher level. This might be because of high local living costs. Or some JG employers might want to offer additional benefits (at their own cost) to workers, including housing allowances for high rent areas.
14. Limited pilot programs that experiment with different models deviating from what is described above might also be considered. For example, a pilot program run by the Federal government, with all participants hired as federal employees, might be tried before the JG is implemented on a national scale.

Findings from the Simulation

Let's turn our attention to the results of the Levy simulation of a nationwide JG program for the USA, which generally follows the formulation outlined above. (Not every prescription was followed, however. See the report for details.) (Wray, et. al. 2018)

For the purposes of our simulation, we assumed that the JG would provide new jobs for approximately 15 million workers at \$15 per hour. The total employment number was obtained by using a generous estimate of unemployment and underemployment at the time of the study (before COVID) and adding to that a projection of the number of existing very low-wage jobs that might disappear if employers had to compete with the JG wage. Our simulation also assumed the program would include a package of benefits worth 25 percent of the wage bill and cover additional costs at 20 percent of the wage bill. The generous wage and benefit package would become standard across the country, as all private and government sector employers would need to match it to retain workers.

The simulation results showed that this would create an additional 4.2 million private sector jobs due to the additional consumption created by JG income. It would boost real GDP by about \$560 billion per year, while inflation would peak at just 0.74 percentage points above baseline projections (falling to just 0.09 percentage points above the baseline by the end of the program's first decade). The Federal government's budget deficit would rise by a maximum of 1.5 percent of GDP over the baseline, falling to an average of 1.13 percent of GDP by the second half of the decade.

State budgets would improve by \$53 billion per year due to higher tax revenues received, boosted by the higher national GDP—even after raising the wages and benefits of all government employees to the program's standards. These estimates do not take into account of many of the cost savings that would be enjoyed by all levels of government, firms, and households as a result of the increase in wages and reduction in unemployment and poverty rates.

We found that the program would have a significant effect on poverty rates, disproportionately benefiting women and minorities. At \$15 per hour, one full-time worker could lift a family of up to five out of poverty; with one full-time and one part-time worker, a family of eight could rise out of poverty. In 2017, nearly 6 million families lived in poverty even with a full-time worker. We found that with one full-time worker per family in the program, 9.5 million children would be lifted out of poverty. The average income gap for the 8 million families living in poverty in 2017 was \$10,505—which is less than what a half-time job in the JG program would pay. While not all of those living in poverty would take advantage of the JG program, it would have the potential to lift most families out of poverty if they did.

As noted, we did not include in our simulation many of the cost savings that would result from full employment. For example, in 2015, the Federal government spent \$104 billion on Food and Nutrition Service programs (including \$74 billion for the Supplemental Nutrition Assistance Program, \$21 billion for child nutrition programs, and \$6 billion for the Special Supplemental Nutrition Program for Women, Infants, and Children), \$17.3 billion on Temporary Assistance for Needy Families, \$50 billion on housing assistance, and \$67 billion on Earned Income Tax Credits. Additionally, total direct spending by states for social services and income maintenance on public welfare was \$505 billion (this does not include spending on health, policing, or corrections). It is conceivable that if we had included all social and economic benefits—including reductions in poverty, indebtedness, crime, and incarceration and improvements to physical and mental health—the impact on the federal government budget might be insignificant (and the positive impact on state budgets would be larger).

On the other hand, the positive impacts we reported resulted despite negative impacts on low-wage firms. For example, we assumed that between 3.1 million and 6.2 million full-time or part-time employed workers would leave their jobs to enter the JG program because their current wages and benefits were too low, and their employers would not match those of the JG program. Even though a significant proportion of employers paying very low wages were assumed to go out of business, net private sector employment still rose by 4.2 million because of better economic performance. Total employment rose by about 19 million—a significant increase—and workers received higher pay and better benefits.

We will also note that the Fair Model is largely based on historical data and mainstream economic theory. In many ways, it is biased against MMT and the JG. It does not adopt the MMT approach to government spending, so deficits and debt can influence interest rates in a manner inconsistent with our approach (thus, rising budget deficits increase interest rates, causing economic growth to slow and government spending on interest to rise). It has a conventional approach to inflation, employment, and wages. It does not use our “reserve army of the employed” buffer stock approach. It does, however, have a good fit to real-world data and is a widely respected model. A simulation cannot tell us precisely what the impact of implementing a real-world JG program will be, but it does inform our understanding. In the case of our simulation, it confirmed our core beliefs that the JG would improve economic performance with little inflationary impact.

CONCLUSION

The main issues examined in this chapter concern the desirability and feasibility of a JG program. The program is desired because (1) a capitalist system does not, and arguably cannot, generate continuous full employment; (2) no civilized and wealthy, society should allow a portion of its population to go without adequate food, clothing, and shelter; and (3) our society places a high value on work as the means through which most individuals should obtain a livelihood. JG policy cannot resolve all social problems. JG cannot replace all transfer spending. Some individuals will not be able to work in a JG program. Some individuals will not be willing to work. However, the JG will ensure that all those willing to work at the JG wage will be able to obtain a job.

Indeed, 'ability' should be defined as broadly as possible to include virtually all those who are willing to work—including people with disabilities that have heretofore kept them without jobs. There is no reason to impose a narrow efficiency standard to ensure productivity above the JG wage. Any production will normally be better than no production;⁹ if one begins with the belief that all deserve basic necessities, then the least fortunate must be supported whether they work or not. In some cases, the net product might be negative from a narrow economic standpoint because supervision, capital investments, and personal services required to put some people to work (for example, to employ people with severe disabilities) could exceed the economic value of output. However, a rich society can afford inefficiencies, and the non-economic benefits of work can offset the economic costs.

The JG program is feasible. The modern government does not face financial constraints because anything for sale in terms of the domestic currency can be had by delivering the currency. Neither taxes nor bonds finance spending, as we have explained in previous chapters. However, the government faces real constraints to the extent that it can purchase only that which is for sale. Involuntarily unemployed labor power is, by definition, for sale.

Finally, we have argued that the JG operates like a buffer stock to stabilize prices. In this chapter, we have outlined how the government stabilizes the price of the commodity used in its buffer stock scheme. Indeed, we can view the use of the JG to achieve zero unemployment as a buffer stock price stabilization scheme. Just as the gold standard generated full employment of gold, a labor standard or JG buffer stock program will generate full employment of labor. Its stabilizing capacity is much greater and preferred, over the stabilization of the money value of gold because capitalism is a system that relies on the use of labor to produce output for sale at money prices.

As we discussed in Chapter 10, the value of money is related to the money wage unit, or to state it another way, it is equal to the amount of labor time it

can purchase. The JG program establishes the base wage (and benefit package) at which anyone can obtain paid work. As long as the JG wage remains fixed, the value of money is steadied (other wages could rise relative to the JG wage, but if the buffer stock of JG workers is effective, that will act as a countervailing force). As we have explained, the JG program acts like a buffer stock—but it is a buffer stock of the most important input to the production process. Not only does labor go into the production of everything, but there are also tremendous social benefits to ensuring full employment. The JG program not only stabilizes wages, prices, and consumption, and aggregate demand, but it is also socially and politically stabilizing. It allows policy decisions to be made in conditions of relative price stability and with full employment. In the current environment with chronically high unemployment, policy that otherwise would be seen as counterproductive (spurring fossil fuel production and punitive tariffs on imports to “save jobs”) would be less appealing if full employment were the normal state.

Further, our understanding of the credit nature of money advanced throughout this book makes it clear that true full employment through a JG is affordable: there is no question about the government issuing its sovereign currency *running out* of money—it cannot happen. That doesn’t mean that the government faces no constraints. The government must carefully consider political opposition, possible exchange rate effects, domestic resource constraints, and possible inflationary impacts. The government must also consider its own capacity to undertake a universal JG program, and how best to structure the program given its capacity and national constraints. Still, we believe that a government-funded JG program is the best—and perhaps the only—way to pursue true full employment with relative price and financial stability.

Finally, the JG program tackles the fundamental problem with our economic system—a monetary production economy: there is no tendency towards full employment. Keynes called this the problem of insufficient aggregate demand that results from the role that money plays as the “rooster,” setting a standard for returns to capital assets that is inconsistent with full employment. Marx put it differently: capitalism requires a reserve army of the unemployed to maintain a rate of exploitation that generates sufficient profits. In addition, we could add the problems of excessive *inequality* (emphasized by Keynes), financial *instability* (emphasized by Minsky), and excessive reliance on *growthism* (emphasized by the UN’s Human Rights Council) that threatens life on planet Earth.

NOTES

1. Go to: <https://documents.un.org/doc/undoc/gen/g24/069/70/pdf/g2406970.pdf>
2. Go to: <https://documents.un.org/doc/undoc/gen/g23/071/64/pdf/g2307164.pdf>

3. Most economists and policymakers do acknowledge the possibility that “NAIRU” (the “non-accelerating inflation rate of unemployment”) can be variable. But, as Gregory Mankiw insists, “Life is full of tradeoffs. Consumers trade off spending today against saving for tomorrow. Congress trades off tax cuts against deficit reduction. And the Federal Reserve trades off inflation against unemployment.” (Mankiw 1997, p. 36)
4. If JG includes benefits—such as health care—that low-wage private sector jobs typically do not include, then the impacts will be larger as employers will have to increase private sector benefits to compete with the JG.
5. See Mitchell (1998) for presentation of JG as a buffer stock program.
6. For all commodities other than labor (except in the case of slavery), it is more accurate to say that they are owned. Thus, if the government operates a buffer stock program for a commodity (say, wheat), anyone with undesired inventory can always sell it to the government at a fixed price. Note we are referring only to government buffer stock schemes; any private scheme is undertaken with the expectation of making a profit, usually by cornering the market to reap gains from rising prices.
7. Under a gold standard, the money price of gold is set above its value in alternative uses; this is to ensure that it remains monetized, rather than being melted for alternative uses. This does not mean, however, that all gold would flow to the government, because hoards are protection against the possibility of currency debasement (raising the money price of gold).
8. Keynesian policies tend to induce hiring *off the top*. As aggregate demand rises, firms attempt to bid employed workers away from other firms and to induce unemployed workers into the labor force. It is not surprising that very high levels of aggregate demand are required before firms will hire the least skilled workers, only after wages of more desirable workers have been bid up. On the other hand, JG attains full employment by hiring *off the bottom*—taking those workers who are not desired by private employers—and then trying to make these workers more desirable to the private sector, as they gain experience and training.
9. As Keynes noted, there is no need to justify all government employment in terms of what is *profitable* employment (Keynes 1980, p. 270). See Wray (1997) and Tcherneva (2020) for further discussion of possible JG jobs.

12. Conclusion: the challenges ahead

In this book, we have offered a heterodox approach to capitalism that emphasizes the role that money plays in its economic system. While money has existed for thousands of years, it was not so essential to economic life. We have analyzed the economy using a monetary theory of production and have argued that capitalism has an internal logic that we can understand and use for the analysis of its dynamics. By contrast, the orthodox neoclassical theory proceeds from a few simple axioms about rational behavior to construct a framework that is supposed to apply to all human societies (and, in some cases, the framework has even been extended to study the behavior of fish and other animals). Importantly, orthodoxy relegates the role of money to what is effectively an unimportant sideshow—lubricating an otherwise well-functioning market.

We have adopted an MMT approach to money; however, we have placed it within the heterodox monetary theory of production. Some proponents of MMT prefer to emphasize its descriptive nature—i.e., it addresses the question of “how does the government really spend?” And, certainly, MMT has examined this in detail. Some also emphasize the logic of MMT: the government must supply its currency through spending before taxpayers can use it to pay taxes. As a point of logic, that is, of course true, and we can find historical examples in which colonial authorities (including in early America and colonial Africa) first imposed a tax obligation to “drive” demand for the currency they would issue. Most MMT proponents adopt Lerner’s functional finance approach, which insists that fiscal policy should be focused on results, not based on some preconceived notion about a desired budgetary outcome. Finally, most MMT proponents also adopt the Job Guarantee proposal as a specific policy to achieve Lerner’s recommendation that fiscal and monetary policy should be focused on achieving full employment with relative price stability.

However, description, logic, and policy must be consistent with an understanding of the social system under analysis. That requires theory. And theory must be specific, as Minsky argued, to the institutions and structure of the economy we examine. Moreover, Thomas Kuhn argued that theories can only be formulated within a paradigm—a view as to how the world works. The clearest examples of paradigm shifts come in the “hard” sciences—the Copernican revolution or Einstein’s general theory of relativity. Kuhn argued

that paradigms support but also constrain scientific research because they determine what questions are the questions to be asked. Normal science advances within a paradigm, but anomalies accumulate that cannot be explained within the paradigm. Eventually, an Einstein comes along with a completely different view of the world; a new paradigm displaces the old, and many of the anomalies can be explained in the new view. Importantly, the old views are permanently banished from science. There is a lot of controversy concerning the adaptation of Kuhn's arguments to the social sciences. Ironically, in economics no wrong notion has ever been permanently dispatched—each bad idea always rises to live once again. Perhaps the best example is the notion that the economy can be guided by the invisible hand of the market—an erroneous claim.

With that caveat, this book takes the position that Marx, Veblen, and Keynes worked within a new paradigm that fundamentally rejects the dominant Neoclassical approach. All three of the main heterodox approaches adopted the monetary theory of production. All were concerned with the institutions that exist in capitalist economies. All developed theories consistent with the monetary approach, and all rejected the fundamental foundations of orthodoxy. Keynes, in particular, saw his *General Theory* as a *revolution* of economic theory on par with Einstein's *General Theory of Relativity* that revolutionized physics. Unfortunately, Keynes was not nearly so successful as Einstein—Minsky used to say that Keynes's revolution was aborted as some of his ideas were incorporated into the Neoclassical Synthesis while the more radical ideas were jettisoned.

In the remainder of this conclusion, we examine the status of capitalism today, the challenges we face, and how a *monetary theory of production* helps us address them.

WHAT IS DIFFERENT ABOUT CAPITALISM?

Capitalism's reliance on the profit motive is the source of many of our problems. Previous economic systems undertook production to provision life—both for the producing classes and for the parasitic classes. Output may have been limited, but not by a lack of money profits. The feudal lord wanting more wine would take more from peasants; if peasants went without, it was because they couldn't produce more as replacement. In capitalism, by contrast, the focus is profits. If more wine can be produced than can be sold at a profit, the excess gets dumped in the river (as Steinbeck memorialized in *The Grapes of Wrath*: "Then the grapes—we can't make good wine. People can't buy good wine. Rip the grapes from the vines, good grapes, rotten grapes, wasp-stung grapes. Press stems, press dirt and rot.").

As Veblen (1904) put it, capitalism is a comprehensive system of control of industry by pecuniary transaction for pecuniary ends—not to produce goods for use, but for sale. The serviceable myth is that capitalism faces an *economic problem*: scarce resources and unlimited wants. (Wray 2022b) It is serviceable in the sense that it justifies unfulfilled wants. The reality is that scarcity results from profit-seeking behavior: in the short run, economic efficiency requires dumping wine in the river, while in the long run it is economically *efficient* to bulldoze vines to maintain profits (Steinbeck: “The works of the roots of the vines, of the trees, must be destroyed to keep up the price, and this is the saddest, bitterest thing of all”). Manufactured scarcity is required.

To sum up, capitalist production takes the form of “money now for more money later.” Those with access to money control the production process and undertake production only if they expect to end up with more money than they started with.

In Chapter 24 of his *General Theory*, Keynes addressed two fundamental flaws of capitalism. The first is that the incentive to produce is usually too weak to employ all who want jobs. The second is that capitalism creates excessive inequality. Let’s quickly examine these two faults identified by Keynes and then look at the third fault, added by Hyman Minsky: financial instability.

In the orthodox approach, capitalism is presented as a “market” economy. With flexible wages and prices, the economy moves to equilibrium at full employment of all resources—and with utility maximization of consumers (who are “sovereign”) and profit maximization of firms (that serve the interests of sovereign consumers). Output is fairly distributed among all factors of production (land, labor, and capital) such that the return of each is equal to its marginal contribution to production. (Wray 2022b; see also Dugger 1989) If money is used, it is neutral—at least in the long run—meaning that equilibrium prices, wages, and interest rates are all in relative terms. There’s little for government to do except in the case of market “failures,” including most importantly the case of “public goods” where consumption is nonrival and/or nonexclusionary. If inequality exists, it is due to differences in marginal productivity. Irrational discrimination would be competed away—as rational agents take advantage of underpriced factors of production, with their demand pushing rewards to equilibrium levels. Capitalism is also efficient: in equilibrium, all scarce resources have a price that will allocate them to the best use possible. This is a comforting view, but it does not capture the reality of capitalism as a *monetary production economy*.

Like other class societies, capitalism requires inequality and thus justification of it. (Wray 2022b) The first claim follows from the social relations of production: the capitalist controls the means of production and the worker must work for wages to survive. If all had equal wealth, there would be no division of the classes required for the capitalist social system. As John Henry argued,

“humans lived most of their collective lives in egalitarian relationships”, hence it is a puzzle that “*the settled habits of thought* promoting inequality are sufficiently accepted by a large enough segment of the population to allow those benefiting from the inequality ... to impose their interest to the detriment of the rest of the population.” (Henry 2017, pp. 250–1) He goes through early justifications, such as natural inferiority (Aristotle), racial superiority, “god’s plan”, and moral degeneracy (Malthus). Later, the ideology of the free market—including free labor—appeared to offer a justification for inequality.

It is not a coincidence that economics as a discipline arose with the need to explain away capitalism’s inequality using marginal productivity theory: if you are poor, it is because you deserve to be so. What is the purpose of the marginal productivity theory? Not only to justify the existing, unequal distribution but also to hinder attempts to alter it. Redistribution would undermine the noble traits of strength of character and effort by punishing the productive and rewarding the unproductive. (Peach and Adkisson 1997) The (in)famous *Cambridge Controversies* have demonstrated the logical errors made by high theory, but these academic debates have had little impact on economic theory or teaching. A simpler argument is sufficient to counter the neoclassical view: in all human societies, production is social. “The production process involves a whole series of social relations: the ownership of the means of production, the organization of the workplace, the relations within the family and between the sexes, the policies of the state, and continuous struggle over control of all these relations.” (Dugger 1996, pp. 39–40) Given the social nature of production, it is not possible to allocate to any individual, even imprecisely, her contribution. (Wray 2022b)

Minsky (1986) adds a third fundamental fault: capitalism is dynamic, and a “run of good times” will generally increase financial fragility to the point that it produces a financial crisis—and poses a risk of collapse into a depression if a debt deflation is triggered. Very briefly, Minsky believed that robust growth encourages risk-taking, most importantly, leveraging up debt ratios. With the acceptance of higher debt loads, more spending can be undertaken—and as we’ve seen, an increase in investment produces more profit at the aggregate level (as the Kalecki equation shows) and lower private sector surpluses (as shown in the sectoral balances). If rosy expectations of profits are not validated at the individual level, defaults on debt occur. These can trigger more defaults because balance sheets are highly interconnected (the creditors of the defaulters have debt of their own). With more defaults, the value of financial (and real) assets is re-evaluated, and holders try to sell them before they depreciate further. The sell-off can spread through the economy as the attempted liquidation of assets leads to what Irving Fisher called a *debt deflation* process. This is what made the Great Depression so *great*.

Minsky recognized that the post-WWII stabilizers put in place could prevent the worst outcome: the “big government’s deficit” could rise, propping up profit flows, while the “big bank” (central bank) could act as lender of last resort—lending against assets at face value (rather than at the depreciating value) to stop a liquidation panic. However, he warned, *stability is destabilizing*: just knowing that government backstops are there encourages even more risk-taking. Every time the government prevents *it* (a debt-deflation-induced depression) from happening again, it encourages more risk-taking. While Minsky began to formulate his *financial instability hypothesis* in the late 1950s, it didn’t become clear how right he was until the 1980s Savings and Loan Crisis, followed by the 2000s Global Financial Crisis. The fragility created by the *financial engineering* of the final decades of the 20th century made capitalism far more unstable. Minsky had argued that there are “57 varieties” of capitalism and named the current version *Money Manager Capitalism*. He argued that it is so dangerous that another depression like that of the 1930s might happen again. (See Tymoigne and Wray 2014 for more discussion.)

These three faults (unemployment, inequality, and instability) follow directly from the foundation of capitalism on monetary production—where the purpose of production is to end up with *more money*. As we have made clear throughout this book, money has existed for thousands of years, playing well-recognized functional roles as a medium of exchange, store of value, and means of payment. Money still plays all of those roles today; however, its importance goes well beyond that: capitalist production begins with money to end with more money. Production will not be undertaken if it is not expected to generate profit in money form. Furthermore, as Keynes insisted, the return on money itself sets a hurdle that the return on production must exceed. That is, holding money can provide a better return than producing commodities—in which case production will not be undertaken at all. Finally, the measure of success that counts is the accumulation of money, not the satisfaction of economic needs. For these reasons, unemployment, inequality, and instability are outcomes that are built into the nature of capitalism.

CHALLENGES TODAY

We live in an era of multiple and interlinked pandemics: climate catastrophe, rising sea levels, perpetual fires, floods of refugees, poverty, homelessness, chronic and long-term unemployment, and rising inequality, racism, fascism, and right-wing autocrats. Naomi Klein (2019) drives this point home: the challenges are interrelated, and we must tackle all of them at once. She links this to the history of capitalism, a system based on exploitation. Klein goes beyond Marx’s theory of the exploitation of labor: capitalism exploits the environment, the family, and “the other”—that is, other people outside capitalism’s

core. The rise of capitalism in the New World was based on exploiting slaves, as well as the vast productive lands of the Americas, and the genocide of the indigenous population.

We are taught that capitalism thrives because it is efficient—but that is at best misleading. In many ways, it is inefficient and externalizes costs. From the beginning, it exploited Africa—forced to nurture a population until it could be kidnapped, enslaved, and put to work in the New World. Capitalism exploits families, paying meager wages to the breadwinner and expecting the family—largely women—to bear the costs of producing the next generation of workers. Society—including the public sector today—must educate, train, discipline, punish, medicate, and incarcerate the population to keep it employable. It must protect capitalists from real and imaginary domestic and foreign threats and bail them out from their excesses—as in the decade after the GFC. And it must deal with the damage capitalism does to our society and our world—that includes the physical and mental health crises caused by opioid and social media addictions.

Long ago, Baumol warned that in some sectors of the economy it is difficult to achieve significant increases in labor productivity—especially the arts and personal services, including childcare and healthcare. These tend to be shed by the for-profit sector and end up in the government sector. More generally, we nationalize the costs that capitalism doesn't want to bear. It exploits the “free” resources: the earth, trees, animals, water, and air—and leaves it to the state to clean up the mess. This is euphemistically called a “market failure,” as if correcting the “failure” would make capitalism work better. In truth, capitalism requires externalization of such costs. From this perspective, capitalism has never been sustainable—it has always required the frontier, the new territories, the colonies, the cheap labor, and natural resources of the developing world. We are increasingly recognizing that we have nearly destroyed the planet in capitalism's quest to externalize costs.

Furthermore, capitalism needs effective demand generated outside its *market system*. As a thoroughly monetized economy, where the purpose of production is to obtain profit in monetary terms, there's a tendency to stagnation in both the short run (recession or depression) and the long run (secular stagnation). External demand is required—either exports or sovereign government (that faces no financial constraints other than those that are self-imposed). Export potential is constrained by several factors, at least partially outside a nation's control—and, as Keynes said, pits “nation against nation,” sometimes leading to war.

A safer solution is spending by the domestic government. However, consumption-type spending (including transfers) violates capitalism's norm that rewards should be based on contribution to production, interfering with the incentive structure said to be required to produce a workforce. Government

spending on investment-type output increases capacity, compounding the problem of insufficient demand. Hence, the spending should be wasteful—the *Keynesian* solution was military plus welfare spending.

Such large-scale government spending required—according to conventional views—vast revenue-generating facilities. While it is true that weaker governments had trouble imposing and enforcing an adequate system of taxation (or a system of collection of fees, fines, and tithes), the problem was one of redemption, not the need for revenue. As we have discussed, the sovereign issuer of the currency cannot run out of the currency but may face resistance from the population in accepting it. If the population needs to redeem the currency in payment of taxes (or fees, fines, and tithes), then the government can spend it.

In ancient Greece or Rome, coercion was sufficient to enable the state to grab part of the surplus—but weaker states had less power. They became participants in markets—using money to purchase output. While authorities had always chosen the money of account (and, as discussed, played a role in setting up markets), their own money liabilities were not always accepted in payment. The state would then need to use private *giro* money—as we saw, Italian bankers helped to finance the spending of late medieval sovereigns.

The expansion of the economy increased the size of the tax base, while the development of republican forms of government legitimized the transfer of purchasing power to the state. As discussed in earlier chapters, in the pre-modern period, Bullionist policies and forced loans (and sales of monopoly charters) were gradually abandoned as the modern state emerged with the development of capitalism. The development of a central bank provided the government with the means to expand purchasing power essentially at will. The examination of the case of England showed how the state integrated the public and private *giros* to create the modern financial system—with private money dominating markets but the state's money sitting at the top of the money hierarchy, mostly used for clearing accounts.

Deficits by the government allow the accumulation of a portion of society's surplus output by the government (with a portion allocated to transfers such as old-age security). Except at full employment, deficits not only increase the nominal value of the surplus but also generate an expansion of the social surplus in real terms (that is, in the form of produced commodities). At the same time, government deficits generate claims on the government that are held by "savers" or by the central bank. Thus, deficits have real and financial impacts on the economy. For this reason, the government can play a role in stabilizing the economy, in addition to its role of providing necessary inputs. The development of the central bank created the possibility of stabilizing the financial system through the provision of reserves and enhanced the state's ability to use countercyclical deficits as another stabilization tool.

Finally, as argued by Heilbroner (1985), the development of capitalism required the gradual creation and enforcement by the state of property rights and the gradual exclusion of the state from markets. However, capitalism requires a state—not only to provide defense and to fulfill the other functions enumerated by Adam Smith and Milton Friedman, but also to provide those necessary inputs that cannot be provided profitably by markets (for example, an educated workforce, transportation systems, communication systems, and so on). Furthermore, the state must serve and protect the interests of the capitalists. As argued earlier, capitalism developed out of slavery, which needed a relatively strong state to protect the slave owners. Defense today includes protection of domestic capitalists against competition by foreigners—as well as protection of domestic interests in other nations.

What all of this means is that the state must play a big role in the modern capitalist economy. Its spending provides an external source of aggregate demand and profits for capitalist production. It absorbs costs that capitalists do not want to bear. It protects people and the environment from capitalism's excesses. It can attenuate the tendencies of capitalism to generate unemployment, inequality, and instability. The government promotes and protects the interests of domestic capitalists against competition from abroad. And domestic governments can work with foreign governments to rise to global challenges—many of which are created by the nature of capitalism.

CONCLUSIONS ON THE NATURE OF MODERN CAPITALISM AND THE WAY FORWARD

As we finish up this book, the *Doomsday Clock* that assesses the risk of nuclear annihilation is closer to midnight than it ever has been. Even barring nuclear war, climate catastrophe threatens extinction for most large mammals—maybe including humans. How did we get here? Maybe capitalism was *too successful*.

Back in 1848, Marx and Engels wrote of the shocking success of capitalism in *The Communist Manifesto*:¹

The bourgeoisie, during its rule of scarce one hundred years, has created more massive and more colossal productive forces than have all preceding generations together. Subjection of Nature's forces to man, machinery, application of chemistry to industry and agriculture, steam navigation, railways, electric telegraphs, clearing of whole continents for cultivation, canalisation of rivers, whole populations conjured out of the ground – what earlier century had even a presentiment that such productive forces slumbered in the lap of social labour?

This wasn't hyperbole. But the world had not seen anything yet—what economist Robert Gordon (2016) labels the *special century*, 1870–1970—when innovation and economic growth really took off—was still a generation away.

And only a fraction of humanity lived in capitalist economies in 1848. Imagine what Marx and Engels would think now—with capitalism dominating across the planet and approaching nine billion people under its spell?

Almost a century ago, Keynes wrote his masterpiece, *Economic Possibilities for our Grandchildren* (1930). In it, he discussed how *slow* economic progress had been prior to 1700, and by comparison, how rapid and disruptive it had become. He claimed that the difference could be attributed to the accumulation of capital and technical inventions. From our perspective, it resulted from the accumulation of money capital that would be directed toward commodity production to increase the mass of money capital—using money to make more money. Neither technical innovation nor the satisfaction of human needs is a primary goal of capitalism—to the degree that life got better for humans, that was an unintentional benefit. To be sure, government, religious institutions, and a huge variety of other social institutions have supplemented the pursuit of profit and tried to counter its abuses.

Marx and Engels proposed socialism as an alternative to capitalism. Keynes, however, thought that capitalism would, in a sense wither away as productive capacity would soon easily satisfy all human wants, and the *money motive* would be relegated to a subject for study by specialists in insanity.

We used to believe that modern capitalism was capable, not merely of maintaining the existing standards of life, but of leading us gradually into an economic paradise where we should be comparatively free from economic cares. Now we doubt whether the business man is leading us to a destination far better than our present place. Regarded as a means he is tolerable; regarded as an end he is not so satisfactory ... When the accumulation of wealth is no longer of high social importance, there will be great changes in the code of morals ... The love of money as a possession—as distinguished from the love of money as a means to the enjoyments and realities of life—will be recognised for what it is, a somewhat disgusting morbidity, one of those semi-criminal, semi-pathological propensities which one hands over with a shudder to the specialists in mental disease.²

Economics would return to its original role as the study of moral philosophy, and Keynes hoped that the goal would be to make human wants *desirable* wants. Veblen was more pessimistic about eliminating the “instincts” (more properly defined as inculcated habits) of *invidious distinction* and *conspicuous consumption*—but by exposing and berating them, he must have been hoping to promote more desirable instincts such as those related to *parental bent* and *workmanship*. In other words, all three of the major heterodox traditions wanted to reduce and limit the scope of “money making” as the driving incentive of our economy.

While many have interpreted Keynes’s dismissal of the *love of money* as referring to the propensity to hoard (which reduces aggregate demand), we

would make a broader indictment of for-profit enterprise. As Keynes put it in Chapter 24 of the *General Theory*, his approach was *moderately conservative*—he argued that private enterprise does not seriously misallocate resources but fails to fully employ them. This seems to be inconsistent with the monetary theory of production, which—as Veblen argued—emphasizes *pecuniary pursuits* above all else. Capitalism *does* misallocate resources, and it destroys capacity by exhausting the earth’s regenerative ability.

However, capitalism is not a monolith—as Minsky said, it comes in 57 *varieties*. He singled out the two decades following WWII as relatively progressive (although we cannot overlook the Cold War, various hot wars, and continuing racial and gender discrimination that allowed inequalities to exist), in large part because the government played a bigger role as a regulator and by taking initiative to provision the population outside the for-profit system. He advocated more reliance on public investment and consumption (supplied and enjoyed publicly) and less private consumption and investment. He also promoted his version of a job guarantee—to ensure full employment and to direct the program’s employment to serve the public purpose. He worried that the current stage of capitalism—*Money Manager Capitalism*—not only increases the potential for a complete collapse into another Great Depression, but also results in the “capital development of the economy” being “ill-done.” (Note, by capital development he meant promoting capacity that would improve people’s lives.) Today it is ever more apparent that the current stage of capitalism threatens human survival.

Keynes distinguished between *enterprise* and *speculation*—the first refers to “forecasting the prospective yield of assets over their whole life,” while the second is “the activity of forecasting the psychology of the market.” (Keynes 1964, p. 158) He warned that while “[s]peculators may do no harm as bubbles on a steady stream of enterprise,” “the position is serious when enterprise becomes the bubble on a whirlwind of speculation.” (op. cit., p. 159) Minsky’s *Money Manager Capitalism* is often characterized as the domination of finance over industry—or, the *financialization* of the economy. This is probably a better way to look at the challenge posed by this stage of capitalism—not speculation but rather the financialization of every aspect of life.

In the USA, college has become financialized, with graduates facing tens, and even hundreds, of thousands of dollars of debt—student loans that have been securitized and sold off to “investors” who had bankruptcy law changed to make it almost impossible to get out of the debt. Likewise, home mortgages are packaged into securities that are then sliced and diced into tranches for credit raters to bless. Even the riskiest mortgages (*subprimes*) can provide Triple A tranches supposedly as safe as US Treasury bonds. Derivatives like *credit default swaps* allow investors to place bets on homeowners losing their homes to foreclosures. The adventurous investor can place bets on life and

death through *life settlements* (purchasing life insurance policies of those expected to die early) and *peasant insurance* (employers take out life insurance on workers that pays out when they die). Rights to bridge tolls and parking meter revenues are securitized. Local governments are induced to forgo issuing debt in the form of simple bonds in favor of debt with complex interest rate and exchange rate derivatives that end up increasing their indebtedness.

Corporations are no longer focused on their for-profit business but instead on driving up stock prices that will reward executives who hold billions of dollars in options. Today's dominant firms are in the so-called tech and social media sectors. But, really, their product is the intimate private data scraped from billions of users and then financialized and marketed to advertisers, grifters, political manipulators, and private and government security agencies.

Why do we have hundreds of trillions of dollars' worth of "investments" that play no obvious role in *promoting the capital development* of the economy? Why has the *capitalism* that accomplished so much over the past couple of hundred years morphed into one that *financializes* all aspects of life—and endangers life itself?

Recall that the *monetary theory of production* focuses on the accumulation of money capital directed towards commodity production that will end up with more money. The constraints on capitalism are both technical (real resources and know-how) and financial (mobilizing enough money capital and putting it to productive—i.e., profitable—use). The problem today is that there is far too much money capital searching for returns. It is increasingly concentrated in the hands of the few, with inequality rising in rich countries and wages for most workers relatively stagnant. With globalized production and competition from low-wage countries, and with the threat of robots and artificial intelligence taking away jobs (and, hence, wage income), the inducement to invest money capital into commodity production is greatly diminished. Only financial bubbles can temporarily relieve the tendency to stagnation.

What is the alternative? Government must play a bigger direct role, both in providing jobs and public consumption. It must play a bigger role in regulating private business—both financial and non-financial. MMT and the heterodox approaches help because they shift the focus from supposed financial constraints to the recognition that both government money and private credit money can be expanded at will. Neither government nor banks can ever run out. Much of government policy over the past decades has been developed on the belief that government faces tight financial constraints and that the private financial system must be unleashed from regulations so that it can provide more funding for investment. But these are false narratives.

This isn't the place for a laundry list of policy proposals. Our purpose is to provide the underlying framework for understanding the nature and potential of state and private credit money. To put it as simply as possible, whatever we

know how to do, we can afford to do—with the caveat that we have already exceeded the sustainable use of our planet's natural resources and ability to absorb our waste. However, it appears that we do have the knowledge required to change our path to one that is sustainable.

And there is no question about affordability. Forget the Debt Clock (that used to stand in Times Square in NYC). Instead, focus on the real issues we face—the twin threats of the Doomsday Clock and Climate Catastrophe Clock that are rapidly ticking toward midnight.

NOTES

1. <https://www.marxists.org/admin/books/manifesto/Manifesto.pdf>
2. John Maynard Keynes (1932), 'Economic Possibilities for our Grandchildren (1930)', in *Essays in Persuasion*, New York: Harcourt Brace, pp. 358–73, https://www.aspeninstitute.org/wp-content/uploads/files/content/upload/Intro_and_Section_I.pdf

Bibliography

- Arestis, Philip and Malcolm Sawyer (1998), 'Prospects for the Single European Currency and some Proposals for a New Maastricht', paper presented at The Fifth Post Keynesian Workshop, Knoxville, Tennessee, 25 June–1 July.
- Bagehot, Walter (1927), *Lombard Street: A description of the money market*, London: John Murray.
- Becker, Gary (1974), 'A theory of marriage: part II', *Journal of Political Economy*, 82(2), pp. 11–26. doi: 10.1086/260287.S2CID222442284.
- Bell, Stephanie (2000), 'Do taxes and bonds finance government spending?', *Journal of Economic Issues*, 34, pp. 603–20.
- Bell, Stephanie and John Henry (2001), 'Hospitality versus exchange: the limits of monetary economies', *Review of Social Economy*, 59(2), pp. 203–26.
- Bell, Stephanie and L. Randall Wray (2002–03), 'Fiscal Effects on Reserves and the Independence of the Fed', *Journal of Post Keynesian Economics*, Winter 2002–2003, 25(2), pp. 263–271.
- Bell, Stephanie, John F. Henry, and L. Randall Wray (2004), 'A Chartalist Critique of John Locke's Theory of Property, Accumulation, and Money: Or, is it Moral to Trade Your Nuts for Gold?', *Review of Social Economy*, 62(1) March, pp. 51–65.
- Bernanke, Ben S. (1981), 'Bankruptcy, Liquidity, and Recession', *American Economic Association Papers and Proceedings*, 71(2), May, pp. 155–9.
- Bernanke, Ben S. (1983), 'Nonmonetary Effects of the Financial Crisis in the Propagation of the Great Depression', *American Economic Review*, 73(3) June, p. 257.
- Blinder, Alan and Joseph Stiglitz (1983), 'Money, Credit Instruments, and Economic Activity', *American Economic Association Papers and Proceedings*, 73(2), May, pp. 297–308.
- Bloch, Marc (1953), 'Toward a Comparative History of European Societies', in Frederic C. Lane and Jelle C. Riemersma (eds), *Enterprise and Secular Change: Readings in Economic History*. Homewood, IL: Richard D. Irwin, Inc., p. 494.
- Blomquist, Thomas W. (1979), 'The Dawn of Banking in an Italian Commune: Thirteenth Century Lucca', in *The Dawn of Modern Banking, Center for Medieval and Renaissance Studies*, University of California, Los Angeles, New Haven and London: Yale University Press, p. 53.
- Boulding, Kenneth (1950), *A Reconstruction of Economics*, New York: John Wiley & Sons.
- Boulding, Kenneth (1966), *Economic Analysis Volume IT: Macroeconomics*, New York: Harper & Row.
- Bougrine, H. and M. Seccareccia (2002), 'Money, Taxes, Public Spending, and the State within a Circuitist Perspective', *International Journal of Political Economy*, 32(3), pp. 58–79.

- Bowen, William G., Richard G. Davis, and David H. Kopf (1962), 'The Distribution of the Debt Burden: A Reply', *The Review of Economics and Statistics*, 44, pp. 98–9.
- Braudel, Fernand (1973), *Capitalism and Material Life 1400–1800*, New York: Harper Colophon Books, Harper & Row.
- Brunner, Karl (1968), 'The Role of Money and Monetary Policy', *Federal Reserve Bank of St. Louis Monthly Review*, 50(7), pp. 8–24.
- Cacy, J.A. (1968), 'Alternative Approaches to the Analysis of the Financial Structure', *Monthly Review*, Federal Reserve Bank of Kansas City, March 1968, p. 3.
- Calomiris, Charles, R., Glenn Hubbard and James H. Stock (1986), 'The Farm Debt Crisis and Public Policy', *Brookings Papers on Economic Activity*, vol. 2, pp. 441–79.
- Cameron, Rondo (ed.) (1967), *Banking in the Early Stages of Industrialization: A Study in Comparative Economic History*, New York: Oxford University Press.
- Card, David and Alan B. Krueger (1994), 'Minimum Wages and Employment: A Case Study of the Fast-Food Industry in New Jersey and Pennsylvania', *The American Economic Review*, September 1994, 84(4), pp. 772–93.
- Chick, V. (1983), *Macroeconomics After Keynes*, London: Philip Allan.
- Chick, V. (1985), 'Keynesians, Monetarists, and Keynes: the End of the Debate or a Beginning?', in Philip Arestis and Thanos Skouras (eds), *Post Keynesian Economic Theory: A Challenge to Neo Classical Economics*, New York: M.E. Sharpe, Armonk, p. 79.
- Cipolla, Carlo M. (1980), *Before The Industrial Revolution: European Society and Economy 1000–1700*, New York: W.W. Norton & Company.
- Cohen, Edward E. (2008), 'Elasticity of the Money Supply at Athens', Chapter 4, in W.V. Harris (ed.), *The Monetary Systems of the Greeks and Romans*, Oxford: Oxford University Press, pp. 66–83.
- Cook, R.M. (1958), 'Speculation on the Origins of Coinage', *Historia*, 7, pp. 257–62.
- Cramp, A.B. (1962), 'Two Views on Money', *Lloyds Bank Review*, July, p. I.
- Cramp, A.B. (1970), 'Does Money Matter?', *Lloyds Bank Review*, October, p. 23.
- Crawford, M. (1970), 'Money and Exchange in the Roman World', *Journal of Roman Studies*, 60, pp. 40–48.
- Dalton, George (1967), 'Primitive Money', in George Dalton (ed.), *Tribal and Peasant Economies*. Austin, TX and London, UK: University of Texas Press, pp. 254–81.
- Dalton, George (1982), 'Barter', *Journal of Economic Issues*, 16(1), pp. 181–90.
- Dantas, Flavia and L. Randall Wray (2022), *Handbook of Economic Stagnation*, London: Academic Press.
- Davidson, Paul (1965), 'Keynes's Finance Motive', *Oxford Economic Papers*, 17(1) March, p. 47.
- Davidson, Paul (1978), *Money and the Real World*, London: Macmillan.
- Davidson, Paul (1986), 'Finance, Funding, Saving, and Investment', *Journal of Post Keynesian Economics*, 9(1), Fall, p. 101.
- Davies, Glyn (1997), *A History of Money: From Ancient Times to the Present Day*, Cardiff: University of Wales Press.
- Davis, Richard G. (1968), 'The Role of the Money Supply in Business Cycles', *Monthly Review*, Federal Reserve Bank of New York, 50(4) April, p. 63.
- Day, John (1987), *The Medieval Market Economy*, New York: Basil Blackwell.
- De Cecco, Marcello (1987), 'Financial Innovations and Monetary Theory', in Marcello De Cecco (ed.), *Changing Money*, Oxford: Basil Blackwell, p. I.
- Desan, Christine (2014), *Making Money: Coin, Currency, and the Coming of Capitalism*, Oxford: Oxford University Press.

- Desmond, Mathew (2019), 'In order to understand the brutality of American capitalism, you have to start on the plantation', August 14, 2019, <https://www.nytimes.com/interactive/2019/08/14/magazine/1619-america-slavery.html>
- Dillard, Dudley (1980), 'A Monetary Theory of Production: Keynes and the Institutionalists', *Journal of Economic Issues*, 14(2) June, p. 255.
- Dobb, Maurice (1963), *Studies in the Development of Capitalism*, New York: International Publishers.
- Dow, Alexander C. and Sheila C. Dow (1989), 'Endogenous Money Creation and Idle Balances', in John Pheby (ed.), *New Directions in Post-Keynesian Economics*, Cheltenham, UK and Northampton, MA, USA: Edward Elgar, p. 147.
- Dow, J.C.R. and I.D. Saville (1988), *A Critique of Monetary Policy: Theory and British Experience*, Oxford: Clarendon Press.
- Dugger, William M. (1989), 'Instituted Process and Enabling Myth: The two faces of the market', *Journal of Economic Issues*, xxiii(2), pp. 607–15; p. 608.
- Dugger, William M. (1996), 'Redefining Economics', in Charles J. Whalen (ed.), *Political Economy for the 21st Century: Contemporary Views on the Trend of Economics*, Armonk, NY: M. E. Sharpe, pp. 31–43.
- Eltis, David (2002), 'The Volume and Structure of the Transatlantic Slave Trade: A Reassessment', *The William and Mary Quarterly*, 58(1), pp. 17–46.
- Fazzari, Steve and Hyman P. Minsky (1984), 'Domestic Monetary Policy: If Not Monetarism, What?', *Journal of Economic Issues*, 18(1) March, p. 101.
- Felkerson, James A. (2011), Working Paper no. 698 December, '\$29,000,000,000,000: A Detailed Look at the Fed's Bailout by Funding Facility and Recipient', Levy Economics Institute.
- Foley, Duncan K. (1982), 'The Value of Money The Value of Labor Power and the Marxian Transformation Problem', *Review of Radical Political Economics*, 14(2), pp. 37–47.
- Foley, Duncan K. (1983), 'On Marx's Theory of Money', *Social Concept*, 1(1), pp. 5–19.
- Foley, Duncan K. (1989), 'Money in Economic Activity', in Eatwell, John, Murray Milgate, and Peter Newman (eds), *The New Palgrave: Money*, New York and London, W.W. Norton, p. 248.
- Foley, Duncan K. (2018), 'The "New Interpretation" after 35 Years', *Review of Radical Political Economics*, 50(3), pp. 559–68.
- Forstater, Mathew (2004), 'Tax Driven Money: Additional evidence from the history of thought, economic history, and economic policy', Working Paper no. 35, Center for Full Employment and Price Stability, August 2004.
- Forstater, Mathew (2005), 'The Capitalist State and Its Economy: Democracy in Socialism', *Research in Political Economy*, 22, pp. 51–65.
- Foster, J. Fagg (1981), 'The Reality of the Present and the Challenge of the Future', *Journal of Economic Issues*, 15(4) December, p. 963–967.
- Foster, Gladys Parker (1986a), 'The Endogeneity of Money and Keynes's General Theory', *Journal of Economic Issues*, 20(4) December, p. 953.
- Foster, Gladys Parker (1986b), 'The Endogeneity of Money and Keynes's General Theory', Dissertation, University of Colorado.
- Fox, David (2016), 'The Enforcement of Nominal Values to Money in the Medieval and Early Modern Common Law', in David Fox And Wolfgang Ernst (eds), *Money In The Western Legal Tradition: Middle Ages to Bretton Woods*, Oxford: Oxford University Press, pp. 203–23.
- Fullarton, John (1969), *On the Regulation of Currencies: Being an Examination of the Principles on which it is Proposed to Restrict, within Certain Limits, the Future*

- Issues on Credit of The Bank of England and of the other Banking Establishments Throughout the Country*, New York: Augustus M. Kelley.
- Fulwiler, Scott (2020), 'When the Interest Rate on the National Debt Is a Policy Variable (and "Printing Money" Does Not Apply)', *Public Budgeting & Finance*, Fall, pp. 72–94.
- Fuson, Robert H. (1992), *The Log of Christopher Columbus Translated by Robert H. Fuson*, Camden Maine: International Marine Publishing.
- Galbraith, James K. (1997), 'Dangerous Metaphor: The Fiction of the Labour Market', Public Policy Brief, Jerome Levy Economics Institute, Annandale-on-Hudson, NY, No. 36.
- Gardiner, Geoffrey W. (2004), 'The Primacy of Trade Debts in the Development of Money', in L. Randall Wray (ed.), *Credit and State Theories of Money: The contributions of A. Mitchell Innes*, Cheltenham, UK and Northampton, MA, USA: Edward Elgar, pp. 128–172.
- Ginsburg, Helen (1983), *Full Employment and Public Policy: The United States and Sweden*, Lexington, MA: Lexington Books.
- Godley, Wynne (1997), 'Curried EMU - the meal that fails to nourish', *Observer*, 31 August, Business, p. 2.
- Goldsmith, Raymond W. (1987), *Premodern Financial Systems: A Historical Comparative Study*, Cambridge: Cambridge University Press.
- Goodhart, Charles A.E. (1989a), *Money, Information and Uncertainty*, Cambridge, MA: The MIT Press.
- Goodhart, Charles A.E. (1989b), 'Central Banking', in John Eatwell, Murray Milgate, and Peter Newman (eds), *The New Palgrave: Money*, New York and London, W.W. Norton, p. 88.
- Goodhart, Charles A.E. (1997), 'One Government, One Money', *Prospect*, March, http://www.prospect-magazine.co.uk/highlights/one_gov_one/index.html, pp. 1–3.
- Goodhart, Charles A.E. (1998), 'The Two Concepts of Money: Implications for the Analysis of Optimal Currency Areas', *European Journal of Political Economy*, 14(3), pp. 407–32.
- Gordon, Robert J. (2016), *The Rise and Fall of American Growth; The U.S. Standard of Living since the Civil War*, Princeton, NJ: Princeton University Press.
- Graeber, David (2005), 'Value', in *Handbook of Economic Anthropology*, James Carrier (ed.), Cheltenham, UK and Northampton, MA, USA: Edward Elgar, pp. 439–53.
- Graeber, David (2006), 'Turning Modes of Production Inside Out: Or, Why Capitalism is a Transformation of Slavery (short version)', *Critique of Anthropology*, 26(1), pp. 61–81.
- Graeber, David (2011), *Debt: The First 5000 Years*, Brooklyn: Melville House Publishing.
- Graeber, David (2018), 'Debt Service and the Origins of Capitalism', <https://www.youtube.com/watch?v=K0t50D4lQrs>
- Gramley, Lyle and Samuel Chase (1965), 'Time Deposits in Monetary Analysis', *Federal Reserve Bulletin*, 51(10) October, p. 1380.
- Grey, Rohan (2019), 'Banking in a Digital Fiat Currency Regime', in *Regulating Blockchain: Techno-Social and Legal Challenges*, edited by Philipp Hacker, Ioannis Lianos, Georgios Dimitropoulos and Stefan Eich, Oxford University Press, pp. 169–180.
- Grierson, Philip (1965), 'Money and Coinage under Charlemagne', in W. Braunsfels (ed.), *Karl der Grosse*, vol. 1, Dusseldorf, pp. 501–36; reprinted as Chapter XVIII in *Dark Age Numismatics*, pp. 530–3.

- Grierson, Philip (1975), *Numismatics*, London: Oxford University Press.
- Grierson, Philip (1977), *The Origins of Money*, London: The Athlone Press.
- Grierson, Philip (1979), *Dark Age Numismatics*, London: Variorum Reprints.
- Grubb, Farley (2015), 'Colonial Virginia's Paper Money Regime, 1755–1774: A Forensic Accounting Reconstruction of the Data', Alfred Lerner College of Business and Economics Working Paper No. 2015–11, Newark, DE: University of Delaware, Alfred Lerner College of Business and Economics.
- Gurley, John G. and E.S. Shaw (1955), 'Financial Aspects of Economic Development', *The American Economic Review*, XLV (4) September, p. 515.
- Gurley, John G. and E.S. Shaw (1956), 'Financial Institutions and Interrelationships', *Journal of Finance*, 11, May, p. 257.
- Gurley, John G. and E.S. Shaw (1960), *Money in a Theory of Finance*, Washington, DC: The Brookings Institution.
- Harris, Laurence (1981), *Monetary Theory*, New York: McGraw-Hill Company.
- Harris, W.V. (2008), 'The Nature of Roman Money', in W.V. Harris (ed.), *The Monetary Systems Of The Greeks And Romans*, Oxford, UK and New York, USA: Oxford University Press, pp. 174–207.
- Harvey, Philip L. (1999), 'Liberal Strategies for Combating Joblessness in the Twentieth Century', *Journal of Economic Issues*, 33(2), pp. 497–504.
- Harvey, Philip L. (2000), 'Combating Joblessness: An analysis of the principal strategies that have influenced the development of American employment and social welfare law during the 20th century', *Berkeley Journal of Employment and Labor Law*, 21(2), pp. 677–758.
- Harvey, Philip L. (2002), 'Human Rights and Economic Policy Discourse: Taking economic and social rights seriously', *Columbia Human Rights Law Review*, 33(2), pp. 364–471.
- Heilbroner, Robert (1985), *The Nature and Logic of Capitalism*, New York and London, W.W. Norton and Company.
- Heilbroner, Robert (1988), 'The Problem of Value', Chapter 5 in *Behind the Veil of Economics: Essays in the Worldly Philosophy*, New York: W.W.Norton & Co, pp. 104–133.
- Heilbroner, Robert and Lester Thurow (1982), *Economics Explained: Everything You Need to Know About How the Economy Works and Where It's Going*, Touchstone, New York: Simon and Schuster.
- Heilbroner, Robert and Peter Bernstein (1989), *The Debt and the Deficit*, New York: W.W. Norton.
- Heinsohn, Gunnar, and Steiger, Otto (1983), 'Private Property, Debts and Interest Or: The Origin of Money and the Rise and Fall of Monetary Economies', *Studi Economici*, 21, p. 3.
- Heinsohn, Gunnar, and Steiger, Otto (1984), 'Marx and Keynes: Private Property and Money', *Monnaie et Production*, 1, p. 37.
- Heinsohn, Gunnar, and Steiger, Otto (1989), 'The Veil of Barter: The Solution to the "Task of Obtaining Representations of an Economy in which Money is Essential"', in J.A. Kregel, (ed), *Inflation and Income Distribution in Capitalist Crisis: Essays in Memory of Sidney Weintraub*, New York: New York University Press.
- Henry, John (2004), 'The social origins of money: the case of Egypt', in L. Randall Wray (ed.), *Credit and State Theories of Money: The contributions of A. Mitchell Innes*, Cheltenham, UK and Northampton, MA, USA: Edward Elgar Publishing, pp. 79–98.

- Henry, John (2017), 'Brutus is an Honorable Man', *Journal of Economic Issues*, LI(2) June, pp. 249–61.
- Hirway, Indira (2006), 'Enhancing Livelihood Security through the National Employment Guarantee Act: Toward effective implementation of the Act', The Levy Economics Institute Working Paper No. 437, January, www.levy.org.
- Holmes, Alan R. (1969), 'Operational Constraints on the Stabilization of Money Supply Growth', in *Controlling Monetary Aggregates, Proceedings of the Monetary Conference*, sponsored by the Federal Reserve Bank of Boston, June.
- Hudson, Michael (2004), 'The archaeology of money: debt versus barter theories of money's origins', in L. Randall Wray (ed.), *Credit and State Theories of Money: The contributions of A. Mitchell Innes*, Cheltenham, UK and Northampton, MA, USA: Edward Elgar Publishing, pp. 99–127.
- Hudson, Michael (forthcoming), 'Origins of Money and Interest: Palatial Credit, not Barter', in *Temples of Enterprise: Creating Economic Order in the Bronze Age Near East*, ISLET – Verlag.
- Hume, David (1898), 'Of the Balance of Trade', in Thomas Hill Green and Thomas Hodge Grose (eds), *Essays Moral, Political and Literary*, Volume I, London: Longmans, Green, and Company, p. 330.
- Ingham, Geoffrey (2004a), 'The Emergence of Capitalist Credit Money', in L. Randall Wray (ed.), *Credit and State Theories of Money: The contributions of A. Mitchell Innes*, Cheltenham, UK and Northampton, MA, USA: Edward Elgar, pp. 173–222.
- Ingham, Geoffrey (2004b), *The Nature of Money*, Malden, Mass: Polity Press.
- Ingham, Geoffrey (2013a), 'O Sacred Hunger of Pernicious Gold! What Bands of Faith can Impious Lucre Hold?', *European Journal of Sociology*, 54(1), pp. 127–146.
- Ingham, Geoffrey (2013b), 'Revisiting the Credit Theory of Money and Trust', in J. Pixley (ed), *New Perspectives on Emotions in Finance*, London: Routledge, pp.121–139.
- Innes, A. Mitchell (1913), 'What is Money', *Banking Law Journal*, May, pp. 377–408.
- Innes, A. Mitchell (1914), 'The Credit Theory of Money', (December–January 1914), *Banking Law Journal*, pp. 151–68.
- Kahn, R.F. (1954), 'Some Notes on Liquidity Preference', *The Manchester School of Economic and Social Studies*, 22(3) September, pp. 229–57.
- Kahn, Lord (1974), Delivered 6 November 1974 to the British Academy on 'Re-reading Keynes', *Proceedings of the British Academy*, published 1975, pp. 361–91.
- Kaldor, N. (1970), 'The New Monetarism', *Lloyds Bank Review*, July, p. I.
- Kaldor, N. (1980), *Essays on Economic Stability and Growth*, London: Duckworth.
- Kaldor, N. (1985), *The Scourge of Monetarism*, 2nd edition, London: Oxford University Press.
- Kaldor, N. and J. Trevithick (1981), 'A Keynesian Perspective on Money', *Lloyds Bank Review*, January, p. I.
- Kalecki, M. (1954), *Theory of Economic Dynamics*, London: George Allen Unwin, Ltd.
- Karekan, John H. (1967), 'Commercial Banks and the Supply of Money: A Market Determined Demand Deposit Rate', *Federal Reserve Bulletin*, 53(10) October, p. 1699.
- Keynes, J.M. ([1930] 1976), *A Treatise on Money, Volumes I and II*, New York: Harcourt, Brace & Company.
- Keynes, J.M. ([1936]1964), *The General Theory of Employment Interest and Money*, New York and London: Harcourt Brace Jovanovich.

- Keynes, J.M. (1937), 'Alternative Theories of the Rate of Interest', *Economic Journal*, 47, June, p. 241.
- Keynes, J.M. (1939), Preface to the French edition of *The General Theory of Employment Interest and Money*.
- Keynes, J.M. (1971a), *The Collected Writings of John Maynard Keynes*, Volume IV, edited by Donald Moggridge, London and Basingstoke: The MacMillan Press, Ltd.
- Keynes, J.M. (1971b), *The Collected Writings*, Volume V, London: MacMillan.
- Keynes, J.M. (1971c), *The Collected Writings*, Volume VI, London: MacMillan.
- Keynes, J.M. (1973), *The Collected Writings of John Maynard Keynes*, Volume XIV: *The General Theory and After: Part II Defence and Development*, edited by Donald Moggridge, London and Basingstoke: Macmillan/Cambridge University Press.
- Keynes, J.M. (1979), *The Collected Writings* Volume XXIX, London: MacMillan.
- Keynes, J.M. (1980), *The Collected Writings of John Maynard Keynes*, Volume XXVI: *Activities 1940–46, Shaping the Post-war World: Employment and Commodities*, edited by Donald Moggridge, London and Basingstoke: Macmillan/Cambridge University Press.
- Keynes, J.M. (1982), *The Collected Writings of John Maynard Keynes*, Volume XXVIII, edited by Donald Moggridge, London and Basingstoke: Macmillan/Cambridge University Press.
- Keynes, J.M. (1983), *The Collected Writings of John Maynard Keynes*, Volume XI: *Economic Articles and Correspondence, Academic*, edited by Donald Moggridge, London and Basingstoke: Macmillan/Cambridge University Press.
- Keynes, J.M. (1987), *The Collected Writings* Volume XIII, paperback Edition, London: MacMillan.
- Klein, Naomi (2019), *On Fire: The (Burning) Case for a Green New Deal*, New York: Simon&Schuster.
- Knapp, George Friedrich ([1924]1973), *The State Theory of Money*, Clifton, NY: Augustus M. Kelley.
- Kraay, C.M. (1964), 'Hoards, Small Change and the Origin of Coinage', *Journal of Hellenic Studies*, 84, pp. 76–91.
- Kregel, Jan A. (1980), 'Expectations and Rationality within a Capitalist Framework', Paper presented at the American Economic Association meetings, Denver, 1980.
- Kregel, Jan A. (1986), 'A Note on Finance, Liquidity, Saving, and Investment', *Journal of Post Keynesian Economics*, 9, (1), Fall, p. 91.
- Kregel, Jan A. (1987), 'Irving Fisher, Great-grandparent of the General Theory: Money, Rate of Return Over Cost and Efficiency of Capital', Mimeo, Johns Hopkins University Center, Bologna Center.
- Kregel, Jan A. (1988), 'The Multiplier and Liquidity Preference: Two Sides of the Theory of Effective Demand', in Alain Barrere (ed.), *The Foundations of Keynesian Analysis: Proceedings of a Conference held at the University of Paris I-Pantheon-Sorbonne*, New York: St. Martin's Press, pp. 31–250.
- Kregel, Jan A. (1998), 'Price Stability and Full Employment as Complements in a New Europe', paper presented at The Fifth Post Keynesian Workshop, Knoxville, Tennessee, 25 June–1 July.
- Lau, Jeffrey Y.F. and John Smithin (2002), 'The Role of Money in Capitalism', *International Journal of Political Economy*, 32(3), pp. 5–22.
- Laudadio, Leonard (1987), 'Commercial Banks: Market Structure and Short-term Interest Rates', *Journal of Post Keynesian Economics*, 9(4), Summer, p. 632.
- Lavoie, Marc (1984), 'The Endogenous Flow of Credit and the Post Keynesian Theory of Money', *Journal of Economic Issues*, 18(3) September, p. 771.

- Lavoie, Marc (1985), 'Credit and Money: The Dynamic Circuit, Overdraft Economics, and Post Keynesian Economics', in Marc Jarsulic (ed.), *Money and Macro Policy*, Boston-Dordrecht-Lancaster: Kluwer-Nijhoff Publishing, p. 63.
- Lavoie, Marc and M. Seccareccia (1988), 'Money, Interest and Rentiers: The Twilight of Rentier Capitalism in Keynes's General Theory', in O.F. Hamouda and John N. Smithin (eds), *Keynes and Public Policy After Fifty Years, Volume II: Theories and Method*, New York: New York University Press, p. 145.
- Lerner, Abba (1943), 'Functional Finance and the Federal Debt', *Social Research*, 10, pp. 38–51.
- Lerner, Abba (1947), 'Money as a Creature of the State', *American Economic Review*, 37(2) May, pp. 312–17.
- Lerner, Abba (1961), 'The Burden of the Debt', *Review of Economics and Statistics*, 43, pp. 139–41.
- Levine, David (1983), 'Two Options for the Theory of Money', *Social Concept*, I(I) May, pp. 20–29.
- Lockhart, P.R. (2019), 'How slavery became America's first big business', Vox, <https://www.vox.com/identities/2019/8/16/20806069/slavery-economy-capitalism-violence-cotton-edward-baptist#>.
- Lopez, Robert S. (1979), 'The Dawn of Medieval Banking', *The Dawn of Modern Banking*, Center for Medieval and Renaissance Studies, University of California, Los Angeles, New Haven and London: Yale University Press, p. I.
- MacDonald, George (1916), *The Evolution of Coinage*, Cambridge and New York: University Press, G.P. Putnam's Sons.
- Maddox, Thomas, Esq. (1969), *The History and Antiquities of the Exchequer of the Kings of England in Two Periods*. Vols. 1 and 2, 2nd ed. New York: Greenwood Press.
- Malinowski, Bronislaw (1921), 'The Primitive Economics of the Trobriand Islanders', *The Economic Journal*, 31(121), pp. 1–16.
- Malinowski, Bronislaw (1932), *Argonauts of the Western Pacific: An Account of Native Enterprise and Adventure in the Archipelagoes of Melanesian New Guinea*, London: George Routledge & Sons, Ltd.
- Mankiw, N.G. and Summers, L. (1984), 'Do Long-Term Interest Rates Overreact to Short-Term interest Rates?', *Brookings Papers on Economic Activity*, No. 1, p. 223.
- Mankiw, Gregory (1997), 'Alan Greenspan's Tradoff', *Fortune*, December 8, p. 36.
- Marx, Karl (1909), *Capital, Volume III*, Chicago: Charles H. Kerr and Company.
- Marx, Karl (1972a), *Capital, Volume III*, London: Lawrence and Wishart.
- Marx, Karl (1972b), *Theories of Surplus Value, Part 3*, London: Lawrence and Wishart.
- Marx, Karl (1978), *Capital, Volume I*, Moscow: Progress Publishers.
- Mcintosh, Marjorie K. (1988), 'Money Lending on the Periphery of London, 1300–1600', *Albion*, 20(4), Winter, pp. 557–71.
- Minsky, Hyman (1957), 'Central Banking and Money Market Changes', *Quarterly Journal of Economics*, 71(2) May, p. 171.
- Minsky, Hyman (1965), 'The Role of Employment Policy', in M.S. Gordon (ed.), *Poverty in America*, San Francisco, CA: Chandler Publishing Company.
- Minsky, Hyman (1975), *John Maynard Keynes*, New York: Columbia University Press.
- Minsky, Hyman (1978), 'The Financial Instability Hypothesis: A Restatement', *Thames Papers in Political Economy*, Thames Polytechnic.
- Minsky, Hyman (1986), *Stabilizing an Unstable Economy*, New Haven, CT: Yale University Press.

- Minsky, Hyman (1992), 'Prices in a Financially Sophisticated Capital-Using Capitalist Economy', Hyman P. Minsky Archive. 35. https://digitalcommons.bard.edu/hm_archive/35
- Minsky, Hyman (2013), *Ending Poverty: Jobs, not welfare*, Annandale-on-Hudson, NY: Levy Economics Institute.
- Mints, Lloyd W. (1945), *A History of Banking Theory in Great Britain and the United States*, Chicago, Illinois: University of Chicago Press.
- Mitchell, William F. (1998), 'The Job Guarantee and the NAIRU', Centre of Full Employment and Equity, Working Paper No. 98–01.
- Mitchell, William F. and L. Randall Wray (2005), 'In Defense of Employer of Last Resort: a response to Malcolm Sawyer', *Journal of Economic Issues*, 39(1), pp. 235–45.
- Mitchell, William F. and Joan Muysken (2008), 'Full employment abandoned: shifting sands and policy failures', Working Paper No. 08–01, Centre of Full Employment and Price Stability.
- Moore, Basil (1984), 'Contemporaneous Reserve Accounting: Can Reserves be Quantity-Constrained?', *Journal of Post Keynesian Economics*, 7(1), FaU, p. 103.
- Moore, Basil (1985), 'Wages, Bank Lending, and the Endogeneity of Credit Money', in Marc Jarsulic (ed.), *Money and Macro Policy*. Boston-Dordrecht-Lancaster: Kluwer-Nijhoff Publishing, p. 1.
- Moore, Basil (1988a), 'The Endogenous Money Supply', *Journal of Post Keynesian Economics*, 10(3), p. 372.
- Moore, Basil (1988b), *Horizontalists and Verticalists: The Macroeconomics of Credit Money*, Cambridge: Cambridge University Press.
- Moore, Basil (1988c), 'Keynes's Treatment of Interest', in O.F. Hamouda and John N. Smithin (eds), *Keynes and Public Policy After Fifty Years*, Volume II: Theories and Method, New York: New York University Press, p. 121.
- Mott, Tracy (1985–86), 'Towards a Post-Keynesian Formulation of Liquidity Preference', *Journal of Post Keynesian Economics*, 8(2), p. 222.
- Munroe, John H. (1979), 'Bullionism and the Bill of Exchange in England, 1272–1663: A Study in Monetary Management and Popular Prejudice', in *The Dawn of Modern Banking*, Center for Medieval and Renaissance Studies, University of California, Los Angeles, New Haven and London: Yale University Press, pp. 169–239.
- Parguez, A. (2002), 'A Monetary Theory of Public Finance', *International Journal of Political Economy*, 32(3), pp. 80–97.
- Papadimitriou, Dimitri and L. Randall Wray (2021), 'Still Flying Blind after All These Years: The Federal Reserve's Continuing Experiments with Unobservables', Public Policy Brief No. 156, December, Levy Economics Institute of Bard College.
- Panico, C. (1982), 'Interest and Profit (a Study in Marx's Writings)', Unpublished Manuscript, March.
- Parguez, Alain and Mario Seccarrecia (2000), 'The credit theory of money: the monetary circuit approach', in John Smithin (ed.), *What is Money?*, London, UK and New York, USA: Routledge, pp. 101–23.
- Peach, James T. and Richard V. Adkisson (1997), 'Enabling Myths and Mexico's Economic Crises (1976–1996)', *Journal of Economic Issues*, xxxi(2) June, pp. 567–74.
- Peacock, Mark S. (2013), 'Accounting for money: The legal presuppositions of money and accounting in ancient Greece', *Business History, Taylor & Francis Journals*, 55(2), pp. 280–301.

- Pesek, B.P. and T.R. Saving (1967), *Money. Wealth and Economic Theory*. New York: MacMillan.
- Polanyi, Karl (1957), *The Great Transformation*, New York: Rinehart & Company, Inc.
- Polanyi, Karl (1971), 'Aristotle Discovers the Economy', in Karl Polanyi, Conrad M. Arensberg, and Harry W. Pearson (eds), *Trade and Market in the Early Empires*, Chicago: Regnery Company, p. 64.
- Prasch, Robert (2003), 'How is Labor Distinct From Broccoli? Some Unique Characteristics of Labor and Their Importance for Economic Analysis and Policy', Middlebury College Working Paper Series 03–30, Middlebury College, Department of Economics.
- Prestwich, Michael (1979), 'Italian Merchants in Late Thirteenth and Early Fourteenth Century England', in *The Dawn of Modern Banking*, Center for Medieval and Renaissance Studies, University of California, Los Angeles, New Haven and London: Yale University Press, p. 77.
- Radcliffe Committee (1959), 'Committee on the Working of the Monetary System', HMSO, London.
- Rawls, J. (1971), *Theory of Justice*, Cambridge, MA: Harvard University Press.
- Redish, Angela (1987), 'Coinage, Development of', in John Eatwell, Murray Millgate and Peter Newman (eds), *The New Palgrave*, New York: W.W. Norton, pp. 376–7.
- Ricardo, D. (1951), *Works and Correspondence*, Volume III, edited by P. Sraffa and M. Dobb, Cambridge: Cambridge University Press.
- Riu, Manuel (1979), 'Banking and Society in Late Medieval and Early Modern Aragon', in *The Dawn of Modern Banking*, Center for Medieval and Renaissance Studies, University of California, Los Angeles, New Haven and London: Yale University Press, p. 131.
- Robinson, Joan (1971), *Economic Heresies*, New York: Basic Books.
- Robinson, Joan (1979), *The Generalisation of the General Theory and Other Essays*, New York: St. Martin's Press.
- Rockman, Seth (2014), Review Essay: 'What Makes the History of Capitalism Newsworthy?', *Journal of the Early Republic*, 34(Fall), pp. 439–67.
- Rousseas, Stephen (1985), 'A Markup Theory of Bank Loan Rates', *Journal of Post Keynesian Economics*, 8(1), Fall, p. 135.
- Rousseas, Stephen (1986), *Post Keynesian Monetary Economics*, Armonk, NY: M.E. Sharpe.
- Samuelson, Paul A. (1973), *Economics*, 9th edition, New York: McGraw-Hill.
- Sayers, R.S. (1957), *Lloyds Bank in the History of English Banking*, Oxford: The Clarendon Press.
- Schmandt-Bessarar, Denise (2014), 'The Evolution of Writing', in James Wright (ed.), *International Encyclopedia Of Social And Behavioral Sciences*, New York: Elsevier.
- Schumpeter, Joseph A. (1934), *The Theory of Economic Development: An Inquiry into Profits, Capital, Credit, Interest and the Business Cycle*, Cambridge, MA: Harvard University Press.
- Schumpeter, Joseph A. (1951), 'The Creative Response in Economic History', in Richard V. Clemence (ed.), *Essays on Economic Topics of J. A. Schumpeter*, Port Washington, NY: Kennikat Press.
- Seaford, Richard (2004), *Money and the early Greek mind: Homer, philosophy, tragedy*. Cambridge: Cambridge University Press.

- Seaford, Richard (2008), 'Money and Tragedy', in W.V. Harris (ed.), *The Monetary Systems Of The Greeks And Romans*, Oxford, UK and New York, USA: Oxford University Press, pp. 49–65.
- Sen, A. (1999), *Development as Freedom*, New York: Alfred A. Knopf.
- Skidelsky, Robert (2009), *Keynes: the return of the master*, New York: Public Affairs.
- Smith, Adam [1776](1937), *The Wealth of Nations*, The Cannan Edition, New York: The Modern Library.
- Sraffa, Piero (1932), 'Dr. Hayek on Money and Capital', *The Economic Journal*, 42(165), pp. 42–53.
- Stanfield, J.R. (1986), *The Economic Thought of Karl Polanyi: Lives and Livelihood*, Houndmills, Basingstoke, Hampshire and London, UK: Macmillan.
- Stiglitz, J.E. and Weiss, A. (1981), 'Credit Rationing in Markets with 'Imperfect Information'', *American Economic Review*, 71(3) June, pp. 393–410.
- Tarullo, D.K. (2017), 'Monetary Policy without a Working Theory of Inflation', Hutchins Center Working Paper No. 33. Washington, DC: The Brookings Institution.
- Taylor, Nick (2008), *American-Made: The Enduring Legacy of the WPA: When FDR Put the Nation to Work*. New York: Bantam Books.
- Tcherneva, Pavlina R. (2020), *The Case for a Job Guarantee*, Cambridge: Polity Press.
- Tcherneva, Pavlina R. and L. Randall Wray (2007), 'Public Employment and Women: The Impact of Argentina's Jefes Program on Female Heads of Poor Households', Working Paper No. 519, The Levy Economics Institute of Bard College, https://www.levyinstitute.org/pubs/wp_519.pdf.
- Terzi, Andrea (1986–87), 'The Independence of Finance from Saving: a Flow-of-Funds Interpretation', *Journal of Post Keynesian Economics*, 9(2), Winter, p. 188.
- Tobin, James (1987), *Essays in Economics Volume I: Macroeconomics*, Cambridge, Massachusetts: The MIT Press.
- Tooke, Thomas (and Newmarch) (1848), *History of Prices and the State of the Circulation From 1792 to 1856*, Volumes III and IV, New York: Adelphi Press.
- Tooke, Thomas (1959), *An Inquiry into the Currency Principle: The Connection of the Currency with Prices and the Expediency of a Separation of Issue from Banking*, The London School of Economics and Political Science, Houghton Street, Aldwych, London.
- Townshend, Hugh (1937), 'Liquidity-Premium and the Theory of Value', *The Economic Journal*, 47(185), pp. 157–169.
- Tymoigne, Eric (2024), 'The Origins of the Platonic Approach to the Analysis of Monetary Systems: Retracing European and Chinese Monetary Thoughts on Chartalism, Nominalism and the Origins of Monetary Systems', Levy Economics Institute Working Paper No. 1058, November 2024, Levy Economics Institute.
- Tymoigne, Eric and L. Randall Wray (2014), *The Rise and Fall of Money Manager Capitalism: Minsky's half century from World War Two to the Great Recession*, London and New York: Routledge.
- Udovitch, Abraham L. (1979), 'Bankers without Banks: Commerce, Banking, and Society in the Islamic World of the Middle Ages', *The Dawn of Modern Banking*, Center for Medieval and Renaissance Studies, University of California, Los Angeles, New Haven and London: Yale University Press, p. 255.
- Veblen, Thorstein Veblen (1904), *The Theory of the Business Enterprise*, New Brunswick, New Jersey: Transaction Books.
- von Reden, S. (2010), *Money in Classical Antiquity. Key Themes in Ancient History*. Cambridge: Cambridge University Press.

- White, Lawrence H. (1984), *Free Banking in Britain: Theory, Experience and Debate 1800–1845*, Cambridge: Cambridge University Press.
- Williams, Eric (1944), *Capitalism and Slavery*, reprint, Chapel Hill, NC: UNC Press.
- Wojnilower, Albert M. (1987), 'Financial Change in the United States', in De Cecco, Marcello (ed), *Changing Money*, Oxford: Basil Blackwell.
- Wolfson, Martin (1986), *Financial crises : understanding the postwar U.S. experience*, Armonk, NY: M.E. Sharpe.
- Wood, Elmer (1939), *English Theories of Central Banking Control 1819–1858*, Cambridge, MA: Harvard University Press.
- Wray, L. Randall (1988), 'Profit Expectations and the Investment-Saving Relation', *Journal of Post Keynesian Economics*, 11(1), Fall, p. 131.
- Wray, L. Randall (1990), *Money and Credit in Capitalist Economies: The Endogenous Money Approach*, Cheltenham UK and Northampton, MA, USA: Edward Elgar.
- Wray, L. Randall (1992), 'Alternative Theories of the Rate of Interest', *Cambridge Journal of Economics*, 16(1), pp. 69–89.
- Wray, L. Randall (1997), 'Government as Employer of Last Resort: Full Employment without Inflation', Working Paper No. 213, Jerome Levy Economics Institute, Annandale-on-Hudson, NY, December.
- Wray, L. Randall (1998), *Understanding Modern Money: the key to full employment and price stability*, Northampton, MA: Edward Elgar.
- Wray, L. Randall (2015), *Modern Money Theory: A primer on macroeconomics for sovereign monetary systems*, New York: Palgrave Macmillan.
- Wray, L. Randall (2018), Policy Note 2018/3, May, 'A Consensus Strategy for a Universal Job Guarantee Program', Levy Economics Institute.
- Wray, L. Randall (2022a), *Making Money Work For Us: How MMT Can Transform America*, Cambridge: Polity Press.
- Wray, L. Randall (2022b), 'The 2022 Veblen-Commons Award Recipient: L. Randall Wray: Wampeters, Foma, and Granfalloon: Enabling Myths and Not-So-Innocent Frauds', *Journal of Economic Issues*, 56(2), pp. 294–313.
- Wray, L. Randall and M. Forstater (2004), 'Full Employment and Economic Justice', in D. Champlin and J. Knoedler (eds.), *The Institutionalist Tradition in Labor Economics*, Armonk, NY: M.E. Sharpe, pp. 253–72.
- Wray, L. Randall (ed.) (2004), *Credit and State Theories of Money: The contributions of A. Mitchell Innes*, Cheltenham, UK and Northampton, MA, USA: Edward Elgar.
- Wray, L. Randall, Flavia Dantas, Scott Fullwiler, Pavlina R. Tcherneva, and Stephanie A. Kelton (2018), *Public Service Employment A Path to Full Employment*, Research Project Reports, April 2018, Levy Economics Institute of Bard College.

Index

- apocentric 74, 76, 77, 90
- assets
 - asset and liability management 130
 - financial assets 45, 126, 129–31, 133, 134, 136, 152, 156, 171, 207, 208, 211, 232, 234, 247, 248, 255, 258, 259
 - real assets 45, 129, 131, 152, 295
- automatic stabilizer
 - budget deficit and 194–5
 - job guarantee and 271
 - taxes and 194–5
- Babylonia 18–20, 27, 81
- Bagehot, W. 55–7, 59–60, 62–4, 247–8
- balance sheet approach 12, 206, 216–17, 221–7
- bank 35, 37, 43–7, 52–64
 - bank assets 62, 126, 128, 137, 183, 229
 - bank deposits 4, 59, 87, 99, 115, 123, 124, 129, 136, 142, 156, 158, 183, 187, 190, 226, 247
 - bank lending 35, 38, 65, 68, 119, 123, 126, 129, 139, 220, 222, 225, 226, 233, 296
 - bank liabilities 35, 37, 45, 52, 55, 56, 59, 61, 62, 64, 71, 82, 87, 91, 92, 98, 101, 105, 124, 127, 128, 135, 141, 144, 154, 165, 166, 180, 182, 187, 188, 190, 210, 215, 226, 247–9
 - bank loans 68, 90, 108, 130, 137, 153, 160, 190
 - bank notes 4, 35, 68, 69, 71, 75–6, 91, 102, 107, 226, 247
 - banks, history of 44–7
- Bank Acts, The 60–62
- bank assets 62, 126, 128, 137, 183, 229
- bank deposits 4, 59, 87, 99, 115, 123, 124, 129, 136, 142, 156, 158, 183, 187, 190, 226, 247
- Banking Principle 99
- Banking School 95–104, 142
- bank lending 35, 38, 65, 68, 119, 123, 126, 129, 139, 220, 222, 225, 226, 233, 296
- bank liabilities 35, 37, 45, 52, 55, 56, 59, 61, 62, 64, 71, 82, 87, 91, 92, 98, 101, 105, 124, 127, 128, 135, 141, 144, 154, 165, 166, 180, 182, 187, 188, 190, 210, 215, 226, 247–9
- bank loans 68, 90, 108, 130, 137, 153, 160, 190
- banknotes 4, 35, 68, 69, 71, 75–6, 91, 102, 107, 226, 247
- Bank of England 53, 55–64, 92, 96, 99, 102, 104, 248
- bank reserves 4, 5, 59, 62, 86, 125, 129–30, 137, 143, 145, 175, 179, 181–5, 188–90, 192, 196–7, 202, 216, 222–4, 226–32, 234, 283
- barter economy 7, 8, 15, 244
- base 130–32
- Bell, S. *see* Kelton, S.
- bills of exchange 18, 36–7, 39, 43–4, 46, 47, 50, 53, 55, 56, 64, 68, 72, 92, 100, 107, 109
- borrowed reserves 46, 211
- buffer stock employment 272
- Bullionism 34, 47–50, 53, 54, 57, 63
- burden of the debt 197, 199–201
- business cycle 163, 271, 274, 283
- Cameron, R. 43, 56, 58, 60, 244, 246, 247, 249
- capital
 - capital, marginal efficiency of 111, 152, 159, 167

- physical capital 105, 252, 258, 259
- real capital 45, 129, 152, 211, 295
- capital accumulation 105, 207, 238, 244–6, 300, 302
- capital flight 99, 117
- capital gains 119–20
- capitalism
 - and accumulation of capital 105, 207, 238, 244–6, 300, 302
 - characteristics of 239
 - and commodification of labor 239
 - and exploitation of labor 254–5, 260, 296
 - and free labor 240, 243, 295
 - and industrial revolution 238, 240
 - logic of xi, xiii, 9, 245–7, 249, 252
 - and middle passage 241
 - and triangular trade 240–42
- cash, cash flow 111–12, 140
- central bank
 - central bank clearing (drain; mechanism) 53, 59, 61, 63, 86, 114, 124, 164, 171, 190, 229
 - central bank control vii, 4, 123, 131, 139
 - central bank, creation of 34, 53–5, 92
 - central bank discount window 124, 136, 138, 220–22
 - digital currency 230–31
 - discount window 124, 136, 138, 220–22
- certificate of deposit 164
- Chartal 73, 74
- Chartalism, Chartalist 66, 67, 74, 76, 86–90, 94, 122, 144, 179
- Chick, V. xi, 154, 161
- Circuit Approach 86, 120, 140–42
- clearing drain 114, 124, 190
- clearing mechanism 45, 59, 164, 173
- climate catastrophe 296, 299, 303
- colonial governor 292
- colonial money 72
- colony, colonies 70–72, 240, 242, 297
- commercial bank 129–30, 134, 206–8, 210, 211–13, 224, 247, 249
- commodification of labor 10, 239
- commodity
 - commodity buffer stock 278, 280–83, 288–91
 - commodity money 3, 4, 18, 34, 37, 68, 83–5, 246, 247, 253
 - practical commodity money 34
 - theoretical commodity money 34
 - commodity production 8, 253, 258, 300, 302
- commodity money 3, 4, 18, 34, 37, 68, 83–5, 246, 247, 253
- Communist Manifesto, The 299
- contemporaneous reserve accounting 90, 181, 185, 206, 228
- Continental 53
- convertibility 35, 37, 52, 57, 60, 71, 76, 77, 84, 101, 104, 137
- co-operative economy 244
- counter-cyclical 265, 282
- COVID viii, 169, 170, 194–195, 204, 234, 265, 270
- Cramp, A.B. 96, 99–100, 102
- creation of money 12, 16, 17, 21, 26, 39, 40, 43, 45, 50, 67, 68, 74, 78, 81, 86–7, 95, 114–15, 117, 120, 127, 129, 132, 133, 140, 142, 146, 167, 176, 188–90, 206, 234, 261
- credit
 - credit inflation 117–19
 - credit money x, xi, 5, 12, 34, 35, 38–42, 49, 50, 54–8, 64, 66, 78–80, 82, 86, 89, 90, 92, 94, 95, 107, 109, 122, 141–5, 150, 167, 238, 244, 249, 302
 - credit rationed 161–5
 - credit rationing 158, 246
 - credit theory of money viii, 78
 - flow of credit 120, 141
- Currency School 96–104, 120
- currency sovereignty *see* sovereign currency
- current account
 - current account balance 202
 - current account deficit 202, 234, 259, 274
 - current account surplus 176, 259
- Davidson, P. 140, 149, 150, 154–5, 209, 255
- dealer banks 184, 196, 224, 236
- debt
 - debt deflation 126–7, 274, 295, 296
 - foreign debt 246

- government debt 31, 32, 45, 55, 59–60, 88, 90, 133–6, 175, 176, 182, 192, 195, 199, 201–3, 205, 207, 227–30, 232–4
- debtors 27, 31, 32, 37, 56, 79–81, 87, 89, 91, 117, 132, 185
- deficit
 - current account deficit *see* current account
 - deficit spending 182, 200
 - government deficit 67, 128, 150, 175–7, 191–2, 194, 200–206, 233, 234, 259, 274, 298
 - trade deficit 62, 97, 99, 103, 118, 191
- deflation 118, 126, 274–5
- deposit
 - demand deposits 4, 46, 52, 87, 98, 125, 126, 135, 141, 165, 180, 190, 206–16, 218–20, 225, 234, 248, 249
 - deposit banking 44, 46, 56, 64, 101
 - deposit multiplier 4, 11, 78, 95, 107, 130, 142, 249
 - time deposits 206–7, 209–11
- depreciation 70–71, 101, 205, 211, 213, 214, 217, 218, 295
- discount window *see* central bank
- discount window
- disintermediation 174
- dot-com bubble 193
- Dow, J.C.R. 163–4
- economic growth 176, 177, 192–5, 203, 204, 206, 234, 267, 299
- Economic Possibilities for Our Grandchildren* 260, 300
- economy
 - barter economy *see* barter economy
 - capitalist economy vii, x, 12, 79, 85, 86, 94, 96, 117, 142, 190, 244–9, 261, 299
 - co-operative economy 244
 - entrepreneur economy 244
 - monetary economy 8, 9, 15–17, 39, 40, 111, 144, 166, 216, 234
- ELR *see* employer of last resort
- employer of last resort 267, 270–72
- endogeneity 86
- endogenous money vii, viii, 12, 86, 94–6, 104–16, 120, 125, 127, 130, 136, 139–45, 148, 149, 156, 157, 159, 161–5, 172, 206
- definition of endogenous money 86
- epicentric 74, 76, 77, 90
- Euro
 - and Draghi “Whatever it takes” 204
 - and ECB 204
 - Economic and Monetary Union (EMU) 201–4
 - and EU Parliament 203
 - euro currency 201–2, 204, 283
 - and fiscal policy 201
 - and Global Financial Crisis 203
 - and government deficit 201, 204
 - and job guarantee 204
 - and Maastricht criteria 201, 204
- Eurodollar 210
- exogeneity 86
- exogenous money 86, 94, 95, 97, 110, 115, 122, 139, 148, 168, 215, 234
- definition of 86
- expectations theory of interest rates 148, 161, 173
- exports 50, 53–4, 92, 154, 191, 197–8, 240, 244, 258, 297
- federal funds rate (fed funds rate) 124, 128, 138, 159–61, 165, 166, 173, 228, 256
- Federal Reserve Bank (Fed), Fed policy 64, 87, 128, 137–8, 156, 159, 161, 164–6, 169–70, 173–4, 180–84, 188, 197, 198, 202, 203, 205, 220, 227–30, 232, 265–6
- feudalism 13, 16, 240
- fiat money 4, 5, 11, 26, 37, 41, 49, 52, 63–6, 76, 78, 82, 84, 85, 93, 98, 100–103, 108, 179, 253, 283
- finance
 - external 132, 246, 247, 256
 - finance motive 112–14, 149, 153, 154, 156, 158, 159
 - finance, Ponzi 139
 - finance, speculative 172
 - finance, state 30, 53–5, 63, 64
 - internal 246
- financial assets 45, 126, 129–31, 133, 134, 136, 152, 156, 171, 207, 208, 211, 232, 234, 247, 248, 255, 258, 259

- financial crisis 63, 167, 203, 267, 295
- financial instability hypothesis 125–9, 142, 151, 152, 166, 267, 296
- financial institution 39–47, 106, 109, 119–20, 125, 127, 129, 133, 136, 150, 155, 156, 158, 171, 172, 174, 226, 232
- financial instruments 35, 40, 45, 87
- financial intermediation 129
- financialization 241, 301
- financialized 301, 302
- financial profiles
 - Ponzi 139
 - speculative 172
- financing, spending 67, 93, 132, 135, 141, 145, 149, 153, 155, 157, 159, 167, 175–8, 182, 186, 187, 197, 202, 206, 225–7, 230, 289
- fiscal policy 66, 67, 86, 177, 180, 187, 201–4, 229, 275, 292
- Foley, D. 105, 253–5, 261, 262, 281
- Forstater, M. 261, 264, 266
- frictional unemployment 266
- Friedman, M. 2, 78, 86, 96, 101, 122, 124, 299
- full employment
 - policy 271
 - sustainability of 267
- Fullwiler, S. 231
- functional finance
 - Abba Lerner's theory of 145, 150, 177–9, 226, 292
 - principles of 145, 178, 179, 189, 201, 204
 - and steering wheel 178
- General Theory (*The General Theory of Employment, Interest, and Money*) 82, 85, 95, 110, 114–16, 121, 150–52, 173, 174, 255, 257, 281, 293, 294, 301
- giro*
 - private 38, 50–51, 63, 76, 85, 90, 145, 179, 181, 230, 298
 - public 85, 92, 145, 236
- Global Financial Crisis (GFC) 2, 138–40, 147, 156, 167, 169–70, 172, 183, 184, 194, 203, 204, 296, 297
- Godley, W. 132, 191, 201, 203
- gold 3–4, 97, 99–100, 102, 108, 117, 118, 125, 281–2, 289
- gold standard 11, 57, 71, 77, 80, 83, 84, 94, 253, 281–3, 289, 291
- Goodhart, C.A.E. 3, 64, 66, 77, 89, 93, 142, 201–3
- government bonds
 - denominated in domestic currency 197–8
 - denominated in foreign currency 198
- government budget balance
 - and deficit 184
 - and surplus 184, 191
- government budget constraint 145
- government debt
 - burden 199–201
 - and debt limit 181–2, 188, 202, 233
 - and default 48, 91, 117, 295
 - and government deficit *see* deficit, government
- Graeber, D. 251, 252, 259–62
- Green New Deal 192
- growth rate
 - and deficit ratio 192–5
 - and government debt 176, 195
- Grubb, F. 71, 72
- Gurley, J.G. 96, 100, 132–6, 142, 155, 156, 190
- Harvey, P. 265, 266
- Henry, J.F. 7, 8, 21, 23, 294–5
- heterodox, theory 2, 15–26, 250, 252, 261, 292, 293, 302
- horizontalism 145
- horizontalist 136–9, 142–3, 145, 148, 157, 161, 165, 169, 171, 173, 249
- how government really spends 176–7
- Hudson, M. xiv, 19–22, 24, 89
- inflation ix, 1, 169–70, 175, 178, 181, 182, 235, 250–51, 266, 267, 271, 274–7, 281–3, 287
- Ingham, G. xiv, 8, 34, 43, 44, 46, 54–58
- Innes, A.M. 24–5, 27–9, 31–4, 42, 66, 77–83, 88, 89, 93, 95, 142–4, 185–6
- innovation, financial innovation 164
- institutionalism, institutionalist xiii, 12, 149, 150, 250
- interest rate

- determination of interest rates 139, 148, 151, 153, 158, 166, 169, 171
 - mark-up theory of determination of 159–60
- and liquidity preference 110, 113, 148, 150–59, 164, 166–8, 171
- long-term interest rate 62, 155, 159–60, 169
- natural interest rate 169–70
- neutral interest rate 256
- short-term interest rate 67, 86, 124, 125, 137, 142, 160, 166, 199
- targeting 126, 128, 131, 138, 139, 145–6, 148, 165–71, 178, 197, 199, 229
- term structure of 124, 172, 231
- theory of 148, 149, 159–61, 173
- investment
 - and interest rate 112–14, 122, 123, 166–7
 - and profits 213, 218–19, 245
 - and saving 112–13, 134, 148–50, 157, 158
- investment bank 247, 248
- involuntary unemployment *see* unemployment
- Jefes y Jefas* (jobs program, Argentina)
 - see* *Plan Jefes y Jefas*
- Job Guarantee (JG)
 - as alternative to growthism 270
 - as alternative to unemployed
 - bufferstock 280, 282, 289
 - as automatic stabilizer 271
 - as a bufferstock 280–83
 - effects on budget deficit 274, 287
 - effects on wages and prices 273–7
 - and minimum wage 272–3, 275–6
 - and poverty 265, 268, 269–70, 287–8
 - and price stabilization 265, 270, 274, 278, 281, 282, 289, 290
 - and reserve army of the employed 277–80
 - and value of money 277, 281, 282, 289–90
 - and wage stabilization 265, 283, 290
- Kahn, R.F. 140, 155, 160, 161, 234
- Kaldor, N. 96, 122–5, 128, 131, 132, 138, 140, 160
- Kalecki, M. 146, 160, 206, 207, 213, 217, 218, 295
- Kelton, S. xiv
- Keynesian
 - economics x, 1, 96, 104, 124, 141, 148, 150, 166, 175, 177, 178, 192, 250
 - policy 282, 291
 - Post Keynesian *see* Post Keynesian
 - pump-priming 177, 267
 - theory 82, 103, 132, 140, 219, 258
 - see also* Neoclassical Synthesis
- Keynes, J.M. vii–viii, x–xiii, 2, 12, 14–16, 18, 19, 26, 34, 52, 65, 66, 77, 79, 82–5, 87–8, 93–6, 108, 110–16, 122, 141, 148–54, 157, 158, 160, 166–8, 172–4, 239, 243, 244, 250, 253, 255–60, 263, 267, 280, 281, 290, 291, 293, 294, 296, 297, 300, 301
- King Charles 92
- Knapp, G.F. vii, viii, 30, 37, 51, 66, 72–6, 79, 81–5, 87, 88, 95, 179, 180
- Kregel, J.A. vii, xiv, 151–3, 204, 255, 256
- labor
 - abstract labor 253, 261
 - dead labor 253, 263
 - free labor 240, 242, 295
 - living labor 253, 263
 - ordinary labor 258, 259, 261, 281
 - unproductive 255
 - value 253–5, 258, 259, 261–3
- labor standard 281, 282, 289
- labor theory of value xiii, 12, 251, 253–5, 257, 258
- lagged reserve accounting 90
- Lavoie, M. 96, 140, 141, 151, 167
- laws of motion 251
- legal tender (laws) 4, 38, 52, 70, 71, 73–5, 77, 81, 82, 84, 85, 88, 144
- lender of last resort 64, 92, 124, 126–8, 158, 235, 248, 296
- liquidity
 - liquidity, demand for *see* liquidity
 - preference
- liquidity preference
 - liquidity preference and demand for money 151, 153–4, 167

- liquidity preference theory of asset prices 12, 255
 - as a theory of value of assets 153, 251–2, 258
- loanable funds, and interest rates 112, 148–51, 175
- loans
 - bank loans 68, 90, 108, 130, 137, 153, 160, 190
 - demand for loans 109, 123, 125, 137, 138, 140, 161, 167
 - loan rate of interest 161, 168, 171, 173
 - loans in Islamic society 38–9
 - loans in Middle Ages 49
- logic of capitalism xi, xiii, 9, 245
- managed money 84–5, 93
- marginal efficiency
 - of capital 111, 152, 167
 - of investment 156, 159, 166, 168
 - of money 156, 159, 167
- mark-up
 - theory of interest rates 159–61, 172, 173
 - theory of prices 160, 250
- Marx, K. x–xiii, 2, 12, 79, 94–6, 104–10, 139, 142, 146, 167, 207, 239, 243, 245, 251–5, 258, 259–61, 263–4, 277, 290, 293, 296, 299, 300
- Mesopotamia 19–20, 23–4
- metalism, metalist 67, 72, 89–91, 93, 94
- Mill, J.S. 96, 97, 100
- mina 18, 21, 23
- Minsky, H.P. ix, vii, x, xiv, 66, 79, 86–8, 93, 96, 125–8, 139, 140, 142, 146, 147, 151, 152, 162, 163, 166, 171, 173, 206, 218, 245, 254, 255, 259, 260, 262, 263, 267, 268, 272, 273, 279, 280, 282, 290, 292–6, 301
- Mitchell, W. 266, 291
- Monetarism 1, 94, 95, 101, 122, 124, 131–3, 173
- Monetarist 94–7, 99, 124, 126, 128, 130–32, 175, 250, 257
- monetary economy 8, 9, 15–17, 39, 40, 144, 166, 216, 234
- monetary equivalent of labor effort 255, 281
- monetary policy 61, 66, 67, 86, 94, 98, 103, 122–5, 144, 167, 169, 170, 180, 198, 201, 227, 250, 270, 275, 292
- monetary system, modern monetary system xi, 34–5, 65, 67, 73, 82, 93, 145
- monetary theory of production x, xi, 85, 116, 250, 292–3, 301–2
- money
 - definition of money xi
 - demand for money *see* money demand
 - endogenous money vii, viii, 12, 86, 94–6, 104–16, 120, 125, 127, 130, 136, 139–45, 148, 149, 156, 157, 159, 161–5, 172, 206
 - exogenous money vii, 86, 94, 95, 97, 110, 115, 122, 139, 148, 168, 215, 234
 - helicopter money 86, 101, 234
 - hierarchy of money 34, 82, 144, 298
 - and inflation 250
 - inside and outside money 132, 134–5
 - and logic of capitalism 238–62
 - money of account xii, 8, 12, 14–16, 20, 21, 23, 25, 26, 34, 37, 43, 51, 52, 55, 65, 66, 79, 81, 83, 85, 88, 89, 95, 110, 144, 179, 238, 261
 - neutral money 4, 134, 142, 156, 249–50, 263, 294
 - origins of money 4–5, 11, 12, 15–26, 144
 - pyramid of money 180, 188
 - quantity of money 86, 96, 97, 100, 102, 104, 106, 108–9, 111, 114, 115, 155, 159, 250, 257, 261
 - supply of money *see* money supply
 - and taxes 71, 88, 90, 177
- money demand 86, 110, 115, 153–4, 157–8, 167, 172, 178
- money financed deficit 175, 182
- Money Manager Capitalism 296, 301
- money now for more money later 154, 294
- money supply 41, 86, 95–7, 99, 102–4, 109, 110, 113–16, 122–6, 128, 130, 131, 133, 134, 136–9, 148, 150, 151, 153–5, 161, 165, 169, 171, 175, 178, 206, 234, 249, 250
- Moore, B. 96, 124–5, 132, 136–40, 142, 150, 151, 159, 160, 165–7, 173, 216

- Mosler, W. xiv
- NAIRU *see* natural rate of unemployment
- National Rural Employment Guarantee Act (jobs program, India) 269
- natural rate of unemployment 170, 291
- Neoclassical Synthesis x, 95, 293
- neoclassical theory 26, 140, 292, 293
- Nersisyan, Y. xiii, xiv
- New Deal 266
- Green 192
- jobs programs
- CCC 268
- WPA 268
- New Monetary Consensus (NMC) 168–70, 250
- New View (Tobin's) 129–32
- off-balance sheet operation vii
- Ohlin, B. 153
- operating procedures (central bank and treasury) 176, 181, 184, 187–9, 192, 225, 226
- overdraft 39, 44, 59, 62, 63, 165, 182–4, 186–8, 223–5, 228, 230–33, 236
- facilities 114, 154, 158
- pandemics x, 194–5, 296
- paradigm
- and anomalies 293
- and revolution of theory 293
- and theory 292–3
- pay community
- private 75, 76, 85, 87, 180
- public 75, 76, 85
- see also giro*
- “pay for” viii, 185–9, 275
- pay token 30, 73
- Peel Banking Act 103
- Plan Jefes y Jefas* 269
- platinum coin (trillion dollar coin) 230
- Post Keynesian
- economics 149–53
- theory x
- of interest rate 148, 151, 172
- of money 116, 140, 250
- price xii–xiii, 20–21, 81, 96–8, 131, 152, 249–61, 274–7, 280–83, 294
- price stabilization 103, 265, 270, 274, 278, 281–2, 289, 290
- pricing 250, 252, 257
- primitive exchange 2, 5, 7
- private property
- and interest vii, 15, 17
- and loans 16, 17
- and origins of money vii, 15–18
- profit
- profit equation 217
- profit expectation 128, 155, 167, 168, 171, 219, 220
- profit rate 120, 229, 253, 254, 258, 261, 262
- profit seeking 7, 166, 229, 246, 294
- Public Service Employment (PSE) 272
- pyramid
- of debt 105, 180
- of liabilities 87, 144
- of money 180, 188
- Quantitative Easing (QE) 139, 167, 169, 171, 183–4, 229, 236
- quantity constraint 104, 119, 130, 162
- quantity theory 95, 103, 123, 124, 132
- Queen Elizabeth 53, 54, 80, 92
- Radcliffe Committee 96, 122–4
- rate of exploitation 254, 255, 277, 278, 290
- rate of profit 166, 211, 229, 253–5, 258, 261, 262
- rate of surplus value 207, 254, 255, 259, 262–3
- recession viii, 123, 124, 139, 164, 165, 173, 175, 192, 195, 203, 204, 250, 267, 274
- redemption
- of bank notes or deposits 35
- in payment of taxes 70–71, 89, 92, 298
- redemption tax 71–2
- reflux, law of 102, 143
- representative money 16, 83
- repurchase agreement 125
- reserves
- bond sales as reserve drain ix, 189
- reserve effects of fiscal policy 143, 188, 189, 231
- reserve requirements 82, 129–30, 137, 138, 144, 173, 209–11, 216, 220

- revolving fund of finance 154, 214, 217
- Ricardo, D. 97, 99, 103, 259
- Robertson, D.H. 153
- Robinson, J. x, 115, 140, 155, 160
- Rousseas, S. 160, 165, 173

- Samuelson, P.A. x, 2–3
- Saville, I.D. 163–4
- saving
 - and investment 112–13, 134, 148–50, 157, 158
 - net nominal saving 283
 - net saving 176, 199, 201, 234
- Sayers, R.S. 59, 61–3
- Schumpeter, J. 73, 78–9, 94, 96, 116–20, 125, 142, 244, 246
- sectoral balance
 - and budget deficit 191
 - and current account balance 191
 - and private surplus 191, 274, 295
- Securities 60, 63, 102, 114, 117, 133, 134, 136, 155, 156, 206–7, 209, 211, 301
- securitization vii, 137, 164, 165, 172, 236
- securitized 165, 301, 302
- Shaw, E.S. 96, 100, 132–6, 142, 155, 156, 190
- slavery
 - and capitalism 240, 242–3, 245, 262, 299
 - and development of modern financial system 240
- sovereign
 - and fiscal policy 180, 187–9, 201–4
 - government viii, 65, 67, 77, 82, 144, 175, 180–87, 198, 224, 227, 297
- sovereign currency viii, 91, 179–83, 186–91, 195–8, 201, 204, 227, 253, 290
- sovereignty, definition of 55
- speculation 160, 301
- stagnation 297, 302
- state money viii, x, 2, 12, 35, 49, 52, 66, 67, 76, 77, 80, 82–5, 89–90, 93, 94, 110, 143–6, 179, 238, 265
- surplus
 - social surplus 239, 245, 298
 - surplus value 104, 207, 254, 255, 259, 262, 263
- sustainability
 - of government debt 192
 - of government deficits 192
- tablets
 - clay xii, 19–20, 24, 26–8, 30, 81
 - cuneiform* 19, 21
 - shubati* 27
- tally, tallies 26–8, 30–33, 55, 80–81
- tax and loan account (TLA) 183–4, 221–5, 236
- tax, taxes
 - taxes-drive-money ix, 90
 - tax liability 70, 71, 88, 90, 190, 221, 222
- Tcherneva, P. xiv, 269, 272, 291
- Tobin, J. 96, 129–30, 132
- Tooke, T. 95–104, 107–8, 121
- treasury 46, 65, 82, 103, 138, 143–6, 175, 180–84, 186–9, 197, 202, 205, 207, 221–5, 228–31, 233, 236
- treasury government account (TGA) 221
- Treatise on Money* vii, 14–15, 82, 95, 110, 114
- two price system (Minsky's) 255, 260
- Tymoigne, E. xiv, 2, 90–93, 296

- uncertainty 8, 16, 17, 48, 70, 107, 135, 161–3, 166, 170, 171, 217–20, 280
- unemployment 166, 170, 178, 193, 195–6, 250, 260, 265, 266, 268, 270–77, 279–83, 287, 289, 290, 296
- natural rate of 170, 291
- United Nation's Human Rights Council 269, 290
- United Nation's Universal Declaration of Human Rights 265
- U.S. exceptionalism 242
- usury laws 36

- value
 - of labor 253–5, 258, 259, 261–3
 - of money xii, xiii, 20, 21, 38, 72, 85, 129, 207, 249–60, 277, 281, 282, 289–90
 - theory xii, 12, 251–2, 258–62
- valuta* money 74–7, 85, 90, 179–80, 188
- velocity 108, 111, 123–6, 129, 131–3, 206, 214, 215, 249
- Volcker, P. 2, 124

- wages, wage bill 214, 233, 234, 244, 250, 255, 259, 287
War on Poverty (The) 268
Wealth 8, 111–12, 132–4, 155, 156, 191, 199, 206, 217, 218
Wergild 12, 15, 22, 23, 26, 40, 77
Williams, E. 240–43, 262
working capital 106, 173, 244–8
Wray, L.R. 2, 8, 83, 141–3, 146, 151, 168, 169, 205, 236, 252, 263, 266, 267, 269, 274, 284, 287, 291, 294–6
Zero Interest Rate Policy (ZIRP) 167, 229