mathematics

Special Issue Reprint

Recent Advances in Business and Industry

Mathematical Analysis, Sustainability Assessment Instruments and Methods

Edited by
Cristina Raluca Gh. Popescu

mdpi.com/journal/mathematics

MDPI

Recent Advances in Business and Industry: Mathematical Analysis, Sustainability Assessment Instruments and Methods

Recent Advances in Business and Industry: Mathematical Analysis, Sustainability Assessment Instruments and Methods

Guest Editor

Cristina Raluca Gh. Popescu

MDPI

Basel • Beijing • Wuhan • Barcelona • Belgrade • Novi Sad • Cluj • Manchester

Guest Editor
Cristina Raluca Gh. Popescu
Department of Economic and
Administrative Sciences,
Faculty of Business and
Administration
University of Bucharest
Bucharest
Romania

Editorial Office
MDPI AG
Grosspeteranlage 5
4052 Basel, Switzerland

This is a reprint of the Special Issue, published open access by the journal *Mathematics* (ISSN 2227-7390), freely accessible at: https://www.mdpi.com/si/mathematics/Recent_Advances_Business_Industry_Sustainability_Assessment_Instruments_Methods.

For citation purposes, cite each article independently as indicated on the article page online and as indicated below:

Lastname, A.A.; Lastname, B.B. Article Title. *Journal Name* **Year**, *Volume Number*, Page Range.

ISBN 978-3-7258-4245-2 (Hbk)
ISBN 978-3-7258-4246-9 (PDF)
https://doi.org/10.3390/books978-3-7258-4246-9

Contents

About the Editor

Cristina Raluca Gh. Popescu

Cristina Raluca Gh. Popescu is Full-Professor Habil. at the University of Bucharest, PhD. Supervisor and Doctoral School Member at The Bucharest University of Economic Studies, Romania, and Award-Winning Scientific Researcher. Having 25+ years of experience, focuses on helping students learn and apply the subjects more effectively, and centers on improving scientific research excellence. She is skilled in Business Administration, Economics, Management, Accounting, and Audit. She has collaborated with The National Institute for Research and Development in Environmental Protection, National Research and Development Institute for Gas Turbines COMOTI, and Ministry of Research and Innovation, Romania. She is a member of the Chamber of Financial Auditors of Romania (CAFR), Romania as Financial Auditor. She has had multiple pivotal roles as the author of 450+ scientific works, with renowned publishers and international coauthors, and is highly cited. She is also an Editor and Editorial Advisory Board Member; Guest, Topic, and Academic Editor; Scientific Committee and Editorial Board Member; Keynote and Invited Speaker; Review Board Member with 600+ reviews. She values a special environment where her in-depth knowledge of creative thinking and academic writing can help deliver accurate research results. She is focused on becoming an outstanding researcher, centered on interdisciplinary topics, in a multicultural environment, constantly driven by passion in promoting the Sustainable Development Goals.

mathematics

Editorial

Preface to the Special Issue on "Recent Advances in Business and Industry: Mathematical Analysis, Sustainability Assessment Instruments and Methods"

Cristina Raluca Gh. Popescu [1,2]

[1] Department of Business Administration, Faculty of Business and Administration, University of Bucharest, 030018 Bucharest, Romania; popescu_cr@yahoo.com or cristina.popescu@man.ase.ro or cristina.popescu@faa.unibuc.ro

[2] Department of Economics and Economic Policy, Economy I Doctoral School, Faculty of Theoretical and Applied Economics, The Bucharest University of Economic Studies, 010374 Bucharest, Romania

MSC: 91G40; 91G70; 91-11; 62H25; 62J05

Recently, it has become pivotal to include mathematics in all domains and crucial to offer a better and more in-depth understanding of mathematics in the forms in which it has been associated with all fields [1]. On the one hand, mathematical analysis offers specialists and profound insights in business and industry concerning the difference between mathematical models and their possible interpretation [2]. On the other hand, sustainability assessment instruments and methods provide specialists in business and industry with evaluation frameworks, as well as the useful tools for identifying major aspects of sustainable development in order to help decision-makers and policy-makers address actions that ought to be taken to make society more sustainable [3].

Moreover, sustainability risk management as a business strategy, these days, takes into consideration the best way in which the profit aims of the companies could be aligned with the environmental, social, and governance policies. As scholars and practitioners have proven on far too many occasions, managing sustainability risks can be successfully achieved by identifying, with the help of mathematical analysis and complex sustainability assessment instruments and methods, the key dimensions that can guide business executives and governance leaders to make an impact that will really make a difference.

Interestingly, theoretical mathematics, also known as research in pure mathematics, is intimately and organically connected with the advancements and developments embodied by applied mathematics.

The Special Issue on "Recent Advances in Business and Industry: Mathematical Analysis, Sustainability Assessment Instruments and Methods" has the immense power to address numerous hot topics of today, among which the following impressive ones could be mentioned: sustainability; green finance; digitalization; digital economy; entrepreneurship; risk management; uncertainty; complex decision processes; finance and financial markets; business processes; Industry 4.0; recession; and, crises, among others.

The paper by Luque González et al. (2021) [4] sheds a new light on sustainability and the way in which it is defined and regarded in different contexts, focusing on creating a highly interesting analysis of 100 theoretical approximations. The authors highlight the construction, evolution, and deployment of the sustainability processes, while displaying the facets of sustainability and their dynamics in ethical, social, and political contexts.

The work of Cai and Guo (2021) [5] addresses finance for the environment, while performing a well-designed scientometrics analysis of green finance. The study brings very important contributions in terms of addressing green finance developments and trends that could motivate both researchers and professionals to have a better view of future studies on making the business process more green.

The focal point of Raut et al. (2021) [6] is on the instruments and methods used to identify indicators of a digital entrepreneurial system. First, the research provides tremendous insights into the positive implications of supporting digital entrepreneurial ecosystems. Second, the work highlights the characteristics of the digital entrepreneurial system of the Republic of Serbia, while making crucial connections with the European Union member states' digital entrepreneurial system and identifies indicators that set back the development of this system, while using appropriate instruments and methods in the process.

The paper authored by Giacomelli et al. (2021) [7] emphasizes the implications of the unsustainability risk of bid bonds in public tenders. The authors present the implications of a forward-looking risk appetite framework, taking into account the case of Italy, while examining the ways in which it "may prevent the demand for unsustainable performance bonds instead of addressing it by rejecting the bidders' requests" [7].

In the paper by Padhi et al. (2021) [8], the authors create a fusion framework for forecasting financial market direction, while using enhanced ensemble models and technical indicators. It is essential to remark on the fact that forecasting the financial market's direction is highly important when aiming to achieve successful trading, yet it is very difficult to understand, in logical terms, all the factors that influence financial markets; it is very challenging to recognize and be able to prevent the responses to market fluctuations and, also, it is overwhelming to design successful trading systems or investment processes.

The research by Pérez-Lechuga et al. (2021) [9] offers a better insight into the area of manufacturing engineering and plant engineering, using the Markov chain theory to model a manufacturing system by process. To illustrate the results derived from their research, the authors provide a numerical example of a real case obtained from a refrigerator factory established in Mexico.

Herman (2022) [10] tackles the challenges as well as the changes in the European Union economy and society in terms of achieving both digital and green transitions. Numerous vital connections are made, in this regard, between the Sustainable Development Goals (SDGs) and digital entrepreneurship.

Nagy and Lăzăroiu (2022) [11] address the Industry 4.0 manufacturing systems, while centering on the implications of globalization, innovation, and digitalization. Likewise, the paper focuses on the integration of computer vision algorithms, remote sensing data fusion techniques, and mapping and navigation tools in the Slovak automotive sector, which provide an interesting and novel approach to the general theme.

The paper by the authors Giacomelli and Passalacqua (2024) [12] aims to bring to light a new internal model approach to measure the Catastrophe Recession Risk, while having central key concepts, such a: credit risk and credit risk models, the European Supervisory Authority, recession, and risk management.

Today, theoretical and applied mathematics in industry and business have not only been a continuous source of motivation for specialists worldwide, but have also been a major source of development in all areas, out of which could be mentioned general management, supply chain management, finance, sales, and marketing, among others [13].

As a Guest Editor of the Special Issue, I am very grateful to all the authors who contributed their articles since, with their effort and work, they created new connections and provided novel insights on vital and pressing issues of today. In line with the aforementioned ideas, I would like to express my gratitude to the reviewers for their precious comments and noteworthy suggestions, which targeted the improvement of the submitted manuscripts, and motivated the authors to bring more arguments to support their work. I strongly believe that this Special Issue had great relevancy and success due to Ms. Krystal Wang, the Section Managing Editor, to whom I would like to express my heartfelt appreciation and gratitude for the on-going collaboration. The goal of the Special Issue on "Recent Advances in Business and Industry: Mathematical Analysis, Sustainability Assessment Instruments and Methods" was to stress the characteristics, importance, and implications of "Dynamical Systems". In addition, the conceptual, theoretical, methodological, empirical, and systematic review studies had, as their ultimate aim, to address the following: the implications of recent advances in business and industry; mathematical analysis; sustainability assessment instruments and methods; inclusive and innovative businesses; numerical and probabilistic techniques applied in sustainability and sustainable economic development and growth; new trends in challenging times (today's pandemic context) and the path to practical statistics in performance and excellence; entrepreneurship; small and medium-sized enterprises; successful business process management; and quality management. Besides the previously mentioned ideas, I hope that all the selected papers for the Special Issue will prove to be impactful for the international scientific community, especially in today's interdisciplinary environment in which it has become a necessity to address solving complex problems with creativity, innovation, and dedicated knowledge in various disciplines and application fields. When closely analyzing the published papers, it is clear that, due to their theoretical and practical features, they are ideal for academics, researchers, scholars, specialists, corporate professionals, and managers.

Funding: This research received no external funding.

Conflicts of Interest: The author declares no conflicts of interest.

References

1. Saravi, M. *Mathematics for Effective Management*; IGI Global: Hershey, PE, USA, 2025. [CrossRef]
2. Shah, N.H.; Mittal, M. (Eds.) *Mathematical Models of Infectious Diseases and Social Issues*; IGI Global: Hershey, PE, USA, 2020. [CrossRef]
3. Christiansen, B.; Shuwaikh, F. (Eds.) *Theoretical and Applied Mathematics in International Business*; IGI Global: Hershey, PE, USA, 2020. [CrossRef]
4. Luque González, A.; Coronado Martín, J.Á.; Vaca-Tapia, A.C.; Rivas, F. How Sustainability Is Defined: An Analysis of 100 Theoretical Approximations. *Mathematics* **2021**, *9*, 1308. [CrossRef]
5. Cai, R.; Guo, J. Finance for the Environment: A Scientometrics Analysis of Green Finance. *Mathematics* **2021**, *9*, 1537. [CrossRef]
6. Raut, J.; Ćelić, Đ.; Dudić, B.; Ćulibrk, J.; Stefanović, D. Instruments and Methods for Identifying Indicators of a Digital Entrepreneurial System. *Mathematics* **2021**, *9*, 2151. [CrossRef]
7. Giacomelli, J.; Passalacqua, L. Unsustainability Risk of Bid Bonds in Public Tenders. *Mathematics* **2021**, *9*, 2385. [CrossRef]
8. Padhi, D.K.; Padhy, N.; Bhoi, A.K.; Shafi, J.; Ijaz, M.F. A Fusion Framework for Forecasting Financial Market Direction Using Enhanced Ensemble Models and Technical Indicators. *Mathematics* **2021**, *9*, 2646. [CrossRef]
9. Pérez-Lechuga, G.; Venegas-Martínez, F.; Martínez-Sánchez, J.F. Mathematical Modeling of Manufacturing Lines with Distribution by Process: A Markov Chain Approach. *Mathematics* **2021**, *9*, 3269. [CrossRef]
10. Herman, E. The Interplay between Digital Entrepreneurship and Sustainable Development in the Context of the EU Digital Economy: A Multivariate Analysis. *Mathematics* **2022**, *10*, 1682. [CrossRef]
11. Nagy, M.; Lăzăroiu, G. Computer Vision Algorithms, Remote Sensing Data Fusion Techniques, and Mapping and Navigation Tools in the Industry 4.0-Based Slovak Automotive Sector. *Mathematics* **2022**, *10*, 3543. [CrossRef]
12. Giacomelli, J.; Passalacqua, L. RecessionRisk^{+}: A Novel Recession Risk Model with Applications to the Solvency II Framework and Recession Crises Forecasting. *Mathematics* **2024**, *12*, 3747. [CrossRef]
13. Popescu, C.R.G. Sustainability Assessment: Does the OECD/G20 Inclusive Framework for BEPS (Base Erosion and Profit Shifting Project) Put an End to Disputes over the Recognition and Measurement of Intellectual Capital? *Sustainability* **2020**, *12*, 10004. [CrossRef]

mathematics

Article

How Sustainability Is Defined: An Analysis of 100 Theoretical Approximations

Arturo Luque González [1,2,*], Jesús Ángel Coronado Martín [3], Ana Cecilia Vaca-Tapia [4] and Francklin Rivas [5]

1 Facultad de Ciencias Humanísticas y Sociales, Universidad Técnica de Manabí, Ave. José María Urbina and Che Guevara, Portoviejo 130105, Ecuador
2 Escuela de Administración-Grupo de Investigación en Dirección y Gerencia, Universidad del Rosario, Calle 12C, 6-25, Bogotá 111711, Colombia
3 Facultad de Educación, Ciencia y Tecnología, Universidad Técnica del Norte, Av. 17 de julio 5-21 y Gral, José María Cordova, Ibarra 100105, Ecuador; jcoronado@utn.edu.ec
4 Facultad de Derecho y Ciencias Sociales, Universidad UTE, Rumipamba y Bourgeois, Quito 170147, Ecuador; ana.c.vaca.tapia@gmail.com
5 Departamento de Informática, Universidad Técnica Federico Santa María, Av. España 1680, Valparaíso 2340000, Chile; firivas@inf.utfsm.cl
* Correspondence: arturo.luque@utm.edu.ec

Abstract: Sustainability processes are imperfect, hence there is a need to analyze their construction, evolution and deployment. To this end, a sample of one hundred sustainability constructs was taken, together with their conceptual approaches, in order to gauge their impact and to ascertain the dimensions to which they belong. A frequency count and categorization were carried out using Google, which saturated in seven dimensions: economic, social, environmental, legal, political, ethical and cultural. A higher-order association of these hierarchies was then proposed, establishing a triad model that indicated only the most representative combinations of dimensions resulting from the extraction of the most significant definitions. From these definitions and in accordance with their frequency of use in Google, it is inferred that the current concept of sustainability is based on the economic-social-ethical category. This highlights the distance between what, a priori, seems to implicitly allow any definition of sustainability and the existing reality.

Keywords: sustainability; google; common benefit; ecological factors; ethics

1. Introduction

This study takes a theoretical and problem-solving approach to the processes of sustainability from an empirical, axiological and propositional point of view through an analysis of the definitions of sustainability. To this end, the conceptual deployment and derived impact of these definitions was studied, including the inherent contradictions and connections arising from their theoretical construction (see Appendix A). Although the exact terminology varies from region to region [1], sustainability has the same underlying significance in all ambits. Today, ecological and social concerns are becoming increasingly important as everyday practices are increasingly regarded as unsustainable [2–5], leading to the emphasis on sustainability as the most effective alternative to the dominant developmental model. This model has been especially questioned and criticized for its role in the global ecological crisis and in the increase in global social inequalities. Common welfare must be achieved in a fair and lasting way through a model of use and management of the environment that recognizes that natural resources are limited and finite [6].

The concept of sustainability can be divided into two parts: regulatory sustainability and positive sustainability. The first consists of the agreements and proposals that were the result of the conceptual framework for sustainable development developed by the United Nations (UN), and the second refers to the scientific analysis of sustainability and sustainable development with an economic and ecological bias [7].

In 1987, the World Commission on Environment and Development (WCED) made public the Brundtland report, better known as "Our Common Future", which specifies that "Humanity has the ability to make development sustainable to ensure that it meets the needs of the present without compromising the ability of future generations to meet their own needs" (p. 23) [8]. From this moment on, the UN made sustainability the principle that governs global development, while, at the same time, academic debate was opened up to a review of the definitions of sustainability by authors, such as Pezzey or Pearce [1].

The concept of sustainability is very broad and varied, depending on the approach used by authors. In a report to the Economic Commission for Latin America (ECLAC) [9], Gallopín characterizes sustainability as an attribute of systems open to interactions with their external world that lack a fixed state of constancy, tending rather toward a dynamic preservation of the essential identity of the system amid continual changes. Salas-Zapata, Ríos-Osorio and Del Castillo [10] describe sustainability as a science by defining it as "the scientific activity that develops around the study of the ability of certain systems to adaptively adjust their socio-ecological relationships to overcome disturbances and maintain some essential attributes and processes". Moreover, the science of sustainability "involves a way of conducting science in a participatory and transdisciplinary way, accepting the uncertainty inherent in real-world phenomena and, consequently, it is more exploratory and comprehensive than predictive in character" (p. 110).

Sustainability can be defined within the limits of the two typologies of weak and strong sustainability [1,7,11]. The first assumes that natural capital and economic capital are fully interchangeable substitutes over a period of time, and the second is based on the economic principle that it is impossible to replace natural resources that pertain to complex natural systems and that, from an economic perspective, are not considered part of the equation. Certainly, in the constitutions of almost all Latin American countries, the development and deployment of sustainability continues to grow despite continuous and systematic environmental degradation and the expulsion and annihilation of cultures [12].

To date, the concept of sustainability remains ambiguous and appears scattered and blurred in the literature [9]. Many of the definitions of sustainability are somewhat abstract and others are decidedly utopian and unquantifiable, hence the objective of this work, which is to analyze these definitions taking into account their cross-disciplinary and dynamic character. From the thorough analysis of academic texts produced by researchers and experts in sustainability processes, 100 theoretical approximations were identified, as set out in Appendix A. Subsequently, a frequency count was carried out using Google, together with a double categorization, in order to discover which dimensions (from the 7 identified) predominate in sustainability processes [13]. From this, a theoretical concept of sustainability was constructed that is a reliable reflection of the current situation, contextualized within the analysis framework used and its intrinsic weaknesses. The Google search was instrumental in ascertaining the real impact of each definition, and it allowed this complex phenomenon to be broken down for analysis. Finally, the necessary factors that contribute to generating sustainability processes were analyzed, together with their institutionalization and regulatory construction. The outcome of the study clarifies whether sustainability processes, by introducing implicit aspects of equality, governance and responsibility, are fully developed in their definitions, or whether the exclusion of certain necessary elements makes them a contradiction of what they claim to be.

2. Theoretical Framework

The National Environmental Policy Act [14] committed the United States to sustainability, declaring it a national policy "to create and maintain conditions under which humans and nature can exist in productive harmony, that permit fulfilling the social, economic and other requirements of present and future generations". While it is true that, at the global level, production and pollution have continued at an unprecedented rate and the instances of ecosocial aggressions are unabated, there is no doubt that this statement of intent was a major step toward addressing many of the global problems that are endogenously created by a model of living based on vanity and continuous consumption. Through this sometimes-involuntary way of life, a large part of the global population is excluded, making them mere subordinates of the developed countries. Sustainability processes are an effective solution to the constant aggressions perpetrated by enterprises, other organizations and individuals on Earth; for example, according to United Nations [15], "today, food losses globally are estimated at 13.8%, equivalent to about $ 400 million." It should be remembered that populations are ecodependent and interdependent and that it follows inexorably that work must be done to achieve zero loss of waste products and, at the exogenous level, allow readjustment through local, national and supranational public policies. For this to happen, lifestyles must be altered in order to minimize (See www.zerowastehome.com and www.trashisfortossers.com, accessed on 1 April 2021) the amount of waste generated and to contribute to the sustainability and conservation of the environment. This requires the development and implementation of the three key concepts of reducing, reusing and recycling [16]. Among the clear possibilities for achieving this are the inclusion of the water footprint (Water use indicator that takes into account both direct and indirect use by a consumer or producer) and environmental footprint (Biophysical sustainability indicator integrating all of the impacts that a human community has on its environment) in any development initiative, thereby bringing to bear fiscal pressure, as well as marshalling public policies; in other words, the greater the corporate and citizen responsibility, the greater the benefits for all. For example, proposals put forward by Indigenous peoples within the Plurinational State of Bolivia are currently gaining strength. The Plurinational Legislative Assembly defines integral development as the continuous process of generating and implementing social, community and citizen measures and actions, and public overseeing, for the creation, provision and strengthening of material, social and spiritual conditions, capacities and media, within the framework of culturally suitable and appropriate practices and actions, which promote relations of solidarity, support and mutual cooperation, and which are complementary and strengthen community and collective ties to achieve Welfare in harmony with Mother Earth [17].

Consumerism has implicitly achieved part of its objectives, that is, to associate the creation of goods and services with the concept of modernity. This movement—despite being a category that refers to the social and historical processes that have their origins in Europe and have been emerging since the Renaissance—has been democratized and trivialized. Today it is becoming more ephemeral and personalized than ever, despite its devastating ecosocial consequences, by proposing that individuals set their goals at will, while the prevailing economic system deploys its influence by characterizing its products and services as both attainable and necessary [18]. Moreover, not all processes are sustainable, nor can many ever become sustainable: the electrification of the economy, for example, does not per se imply sustainability [19,20].

3. Analysis and Discussion

3.1. Delimitation of the Dimensions of Sustainability Processes

The selection and classification of the specified dimensions were designed to reflect the depth, deployment and interconnections observed in the object of study as well as the endogenous characteristics of the concept itself. There are multiple typologies within sustainability processes due to the emergence of new ways of living and producing, as well as the rights associated with these (as claimed by a large part of the public), which are becoming something similar to a social function. The most important consideration was the grouping of related concepts into the seven dimensions: economic, legal, political, social, cultural, ethical and environmental, as set out in Table 1 [13,21,22]. Below is an explanation of the conceptualization of each of these dimensions.

Table 1. Dimensions.

Dimension	Codification of the Definition	Example of the Search String
Economic	Economic or financial aspects affecting and promoting sustainability processes	'new models of economic sustainability' 'speculation' 'maximization of profit' 'disruptive processes' 'social economy'
Legal	Regulations, attitudes and ideologies that produce certainty and contribute to the development of biological systems that remain productive over time	'increase in the powers of bodies responsible for prevention/monitoring/sanctioning to halt regulatory abuse'
Political	Areas that are concerned with power relations, political leadership behaviors, public opinion, international relations and armed conflicts	'promoting equal rights in order to live in an area or neighborhood in a non-invasive way' 'national and transnational regulation'
Social	Aspects that determine and characterize the quality of life of the inhabitants of a given area	'protection according to needs: equity' 'social inclusion' 'protection of a healthy and dignified life (welfare) above other purely economic elements' 'management of unlimited consumption on a finite planet'
Cultural	System of values, beliefs and ways of establishing a society as a mental construct based on geographical identity	'conservation of popular traditions' 'identity' 'cultural heritage'
Ethical	That which shapes man's moral principles and virtues in regard to their responsibility for their actions, to the projection to the community, and to the construction of a healthy coexistence	'values' 'set of norms and customs that direct or value human behavior in a community' 'citizen awareness' 'promotion of social ethics' 'primacy of the person: how we ought to live' 'processes of social responsibility'
Environmental	Natural or altered system in which humanity lives, with all associated social and biophysical aspects and the relationships between them	'pollution' 'reduce, reuse, recycle' 'significant alteration of natural and transformed systems and their resources, caused by human actions' 'contribution to the protection and maintenance of biodiversity'

The economic dimension, despite including a broad social component, suffers from dysfunctions, such as high levels of public and private debt; risks, such as inequality, whereby growth does not often reach the people who need it most; and the deployment and standardization of processes of financial speculation. These risks should be minimized in order to achieve a dignified, fulfilling and harmonious life. There are dangers such as (1) increased trade disputes, (2) stress in financial markets, (3) rising geopolitical tensions, (4) market volatility (in many cases due to financial speculation) and (5) monetary policy adjustment. Hence, there is a need to be aware of changes in the methods of production and their relationship with the environment at all levels by establishing new parameters for production as well as by providing support for public administrations. The social economy attempts to correct many of these dysfunctions, drawing strength from values such as mutual aid, responsibility, democratic values, processes of equality, equity and solidarity, all of which can be supported by binding legal elements [23]. According to Rostow [24] what the economy demands, the law must provide, and the work of the economist is becoming exclusively political through the means provided by legislation, hence the need to establish instruments of verification in order to be able to analyze and redress the harm caused by certain measures that, while legal, may be immoral [25]. According to Leguizamón it will be the responsibility of economic law, because of its great resources and its enormous regulatory power that transcends all fields of social activities, to play a role of the greatest importance in obtaining the objectives of its fundamental principles, which are in perfect interaction with the development of environmental policies (p. 339) [26].

The legal dimension pivots around its relation to the development and deployment of regulations that, in line with the existing historical, social and political circumstances, is legally valid for a given population and thus guarantees the general interest. It is based on the formation of public policies that have the capacity to generate, endogenously, greater ethical, social and responsible practices and behaviors. The establishment of sustainability can be interfered with by such processes as liberalization, deregulation, regulatory hybridization, the formation of lobbies, economic and fiscal reforms, free trade and investment treaties, private arbitration courts and processes that exert legislative pressure [27].

Thirdly, the political dimension consists in the response of civil society and its organization through trade unions, associations, communities, neighborhoods and political parties, among others. One of the tools available to society is civil disobedience as a democratic instrument and, indeed, violence, which must be deconstructed to be analyzed as a response to conditions of despair and constant asymmetries, such as the production of pollution and political and social regulation that favors the minority that holds power and controls capital.

The social dimension is concerned with the cohesion and stability of the population. It relates to the impact of an organization's activities on the social systems in which it operates, including decent employment, human rights, health and social security, public services, population displacement, education and the formation of society based on criteria of public welfare. This also extends to the acceptance of responsibility for the goods and services offered that affect society as a whole, in particular those aspects that determine and characterize the quality of life of local populations at the national level, or through supranational interventions in which the state has the responsibility and capacity to improve the conditions of different social groups [28].

The cultural dimension comprises the need to take an integrated view of development, in which culture is recognized as the multidisciplinary and dynamizing axis of social participation. It accommodates all of the distinctive features that characterize a society or a given social group, including traditions and spiritual, material, intellectual and affective aspects. It also encompasses the arts and literature, ways of life, fundamental human rights, and systems of values and beliefs that have many interrelated meanings [29]. Expressions of the cultural dimension are grouped implicitly and distinguished from others by their reference to humanity.

The ethical dimension includes the fundamental question of how we should we live. Consequently, there is a clear need to direct societies toward rational growth within the natural boundaries set by the Earth. The continuous and systematic deterioration of the environment, the loss of biodiversity, soil degradation and the pollution of surface and groundwater resources represent the continual harm done to all types of ecosystems in favor of the maximization of capital. This has inevitably triggered problems, such as negative climate change, that are often beyond redress and repair.

The environmental dimension highlights the lack of harmonious development between growth and its relationship with the environment since the current developmental model, based in its entirety on the infinite expansion of goods and services, is in conflict with the finite limits of the Earth [30]. It should be noted that the planet is facing a triple environmental emergency of climate change, pollution and loss of biodiversity, hence the need to incorporate the concept of climate neutrality as a strategy in all coordinated actions between stakeholders. It is necessary to incorporate verifiable mechanisms for addressing environmental issues and to provide sufficient resources and tools for their effective deployment. Mechanisms include elements designed to redress, sanction and prevent regulatory abuses, such as those free trade agreements that prioritize the economic over the environmental interest and thereby maintain the existing hegemonic logic that economic growth is facilitated by the abuse of the external physical, chemical and biological components with which living beings interact [31]. The risks to which the Earth is exposed are manifest since the global community has failed to meet its commitments to limit environmental damage by reducing greenhouse gas emissions or putting in place carbon offsets. It is necessary to replace the existing extractivist and anthropocentric vision by promoting the establishment of shared interests that protect all forms of life over private interests dominated by capital. The commodification of nature must be called into question and much of its operating rules redrafted, having been imposed by a global society interconnected directly with the spending capacity of each individual. This phenomenon has been hitherto supported by (un)sustainable companies, governments and supranational organizations. Socioenvironmental tools must be established to break the dynamics of concentration and exclusion based on a clear understanding that the material basis of all human activity is found in nature and the assumption that ecosystems and natural resources have certain limits. Due to their inherent sustainability, environmental issues inexorably have an impact on the work of safeguarding life; therefore, the environmental dimension must be incorporated as a vehicle for development that promotes the future of the entire world population and its quality of life.

3.2. Compilation of Definitions of Sustainability

Sustainability processes have been widely studied and are not limited to one type nor are they linear in nature. Hence, there is a need to begin the experimental procedure with the search for concepts, academic constructs and related elements with the aim of being able to analyze their implications, connections to other fields and constitutive components. This search was based on the academic literature (i.e., research) but also included primary information derived from companies and other organizations, nonacademic experts and other institutional sources. From this starting point, Google Scholar was used to perform Boolean searches of keywords or phrases, such as those in Table 1, which form the basis of the mathematical sets and the logic needed for database searches. Basic operators such as AND, OR and NOT connect the words entered into the search engine to narrow or extend results.

The pre-established definition of the search chains, concomitant to the literature review, saturated in seven categories of sustainability. The process was carried out between 16 November 2020 and 11 February 2021 and resulted in the compilation of 100 units of analysis (UA) (Appendix A), a sample that, according to similar studies, is of sufficient size for a qualitative study (The number of times a particular citation is published is conditioned by the period in which that frequency of occurrence is measured. We assumed that the absolute frequencies thus obtained can only grow or remain stationary temporarily while others increase, resulting in a decrease in the relative frequency. The validity of the analysis is subject to the period in which it is determined to measure how much a certain construct is represented, if at all. Far from being an impediment, we think that interest in the dynamics of sustainability processes lies precisely in this point) [13]. After the Google frequency count of the number of times each of the 100 definitions was cited or referenced, they were assigned dimensions according to the pre-established system of encoding, and the corresponding absolute frequencies, f_i, were obtained as shown in Appendix A. This stage involved several outline conditions for UAs, namely (1) that at least one academic reference be required, (2) that it not be duplicated in different databases and (3) that it may appear in one or more categories, up to a maximum of seven.

3.3. The Study of Dimensions

Once the qualitative approach of this research had been developed, the next step was to explore the scope of each category, as well as its contribution to sustainability processes. The 100 UAs were assigned dimensions within one or more categories, up to a maximum of seven, and the number of times that concept of sustainability had been used in a given time period of approximately 4 months was counted (f_i). In addition, the existence of the seven dimensions unfolds the absolute frequencies in the corresponding f_{kj}, or, to express it another way, the f_i of a unit of analysis k, associated with a category j. Thus, it was possible to obtain the relative weighting of each j dimension according to Equations (1) and (2), respectively, (j = 1, . . . , 7):

$$CD_j = \sum_{k=1}^{y} f_{kj} \tag{1}$$

where CD_j is the scope of each j dimension, and f_{kj} is the frequency of occurrence in Google of each unit of analysis k associated with a j dimension.

$$\%\ PD_j = \frac{CD_j}{\sum_{i=1}^{x} f_i} 100 \tag{2}$$

where PD_j is the relative weighting of each j dimension, expressed as a percentage (%), and f_i is the absolute frequency of occurrence in Google of each UA or bibliographic reference k in a given period of time.

The values obtained from Equations (1) and (2) are given in Table 2. The rating of the social dimension CD_3 corresponds to the sum of all f_{kj} up to the total of y concepts categorized within it. Its contribution to the study is calculated as the relationship between this value and the summation of all absolute frequencies (% PD_j). It follows that the ethical dimension contributes most to sustainability processes, with a factor of 0.69, while the legal and cultural categories are the least representative with a factor of 0.06 and 0.09, respectively.

Table 2. Scope dimension.

j	Dimensions	CD_i	$\%PD_i$
1	Economic	8988	31
2	Legal	985	3
3	Political	7599	26
4	Social	20,975	72
5	Cultural	1291	4
6	Ethical	23,378	80
7	Environmental	7349	25

It should be noted that the rating obtained for each dimension and its relative weighting are directly proportionate, and it should be specified that each CD_j is determined by the use of a citation k by a user, presuming that its use depends on the independent variable of time (set for the seven categories). The contrast is striking between the ethical dimension, found in 80% of definitions, and the legal dimension, which accounts for around one-eleventh of the total. This was a result of using two large sets of variables—CD_j and dimensions—and the fact that, to facilitate the study, the established dimensions were kept constant. However, the rating assigned to each of them might vary, even in those from the same author (Dispersion in the category allocation process was resolved by an iterative technique until the degree of disagreement was below 15% (three steps)). To minimize the impact of qualitative research on data processing, categorization consisted of a three-stage iterative process that included the concept but not the author(s). Given the large volume of definitions used here, this procedure ensured accuracy in the CD_j more reliably than that which would have assigned, in a single stage, a high value to each CD_j (thereby jointly increasing relative weightings), as other authors in the past have done for CS(i)R processes [13]. It should also be mentioned that the 100 concepts were categorized with a minimum of one and a maximum of seven dimensions (Table 3).

Table 3. Number of higher order UA.

Number of Codified Dimensions	Number of Higher Order UA	Weighting in the Google Count (%)	Distribution
7	2	0.6	
6	1	0.3	
5	2	2.7	
4	12	4.8	
3	34	37.4	
2	33	36.2	
1	16	18.1	
	Total = 100	Total = 100%	

3.4. Triad Model and Combination of Dimensions

By exploring Appendix A, it may be observed that the sustainability concepts studied constitute anything from one to several categories with an associated f_{kj}; the volume of information generated is compiled in Table 3. Here, the number of dimensions included in a given definition, the number of definitions that belong to it and the weighting fraction of the count performed in Google may all be inferred. Table 3 shows that, on average, just over 70% of the sustainability concepts studied in this article have been cited, using only dyads and triads, with triads having greater weighting in the Google count.

Table 4 shows the different types of triads identified in the search for each set of definitions (through a combination of dimensions resulting in 100 UAs), as well as their weighting fraction from 37.4% of the Google count. The results show that the association by triads of 34 units of analysis is that which carries the greatest weighting. Thus, a model of triads is defined consisting of the breakdown, with respect to the initial sample, of the 34 UAs catalogued for the corresponding study of their individual contributions (% weighting fraction) as a triad of sustainability dynamics (Table 4).

Table 4. Triad model.

Triad Number	Triads	Number of UA Included	Weighting Fraction (%)
6	Economic-Political-Social	2	0.6
11	Economic-Social-Ethical	8	6.7
12	Economic-Social-Environmental	7	6.7
13	Economic-Cultural-Ethical	1	0.1
15	Economic-Ethical-Environmental	4	6.2
27	Political-Social-Ethical	4	44.6
32	Social-Cultural-Ethical	1	0.2
33	Social-Cultural-Environmental	1	0.5
34	Social-Ethical-Environmental	6	34.4
		Total = 34	Total = 100%

The data obtained from Table 4 suggest that sustainability processes are vectorized by a political, social and ethical category, while the dimensions that are weakly defined are the economic (for example, triads 6 and 13, with 0.6% and 1%) and the cultural (for example, triads 13 and 32 and 33, with 0.1%, 0.2% and 0.5%), while the legal category does not figure at all.

If we apply combinatorial analysis, the grouping of seven elements by threes provides 35 possible triads, such as those shown in Table 5. Of these, only nine have been shown in the study, excluding the remaining 26. This information was collected by taking the individual contribution of each dimension or % PD_j and taking the algebraic sum of the triad to obtain its relative importance, in decreasing order (Table 5).

At the top of the table, labelled "appearing", the data are essentially the same as those already obtained in Table 4 and do not provide additional information, except by the order of their relevance score. The interval of relative importance 21, 8, 31, 28, 4, 29, 18, 35, 25, 10, 2, 26 and 16 may be defined in which those triads not featuring in the categorization had an equal probability of having done so.

The objective is to try to demonstrate whether, in the existing sustainability processes contemplated by experts, there are one or more dimensions in addition to the established triad 11. To achieve this, the entire contribution of the economic, legal and successive dimensions was summed, as shown in Table 6. This shows the most representative dimensions in both sections since, within the "not appearing" group, they represent a mirrored value or reflection of their greater contribution to sustainability; in addition, within this range, the legal category appears as a "lost" dimension within the current processes of sustainability. Table 6 also shows that the cultural category vectorizes sustainability processes only residually.

Table 5. Relative importance.

	Triad	Economic	Legal	Political	Social	Cultural	Ethical	Environmental	Relative Importance (0/1)
Appearing	11	0.31			0.72		0.80		1.83
	27			0.26	0.72		0.80		1.78
	34				0.72		0.80	0.25	1.77
	32				0.72	0.04	0.80		1.56
	15	0.31					0.80	0.25	1.36
	6	0.31		0.26	0.72				1.29
	12	0.31			0.72			0.25	1.28
	13	0.31				0.04	0.80		1.15
	33				0.72	0.04		0.25	1.01
Not appearing	21		0.03		0.72		0.80		1.55
	8	0.31		0.26			0.80		1.37
	31			0.26			0.80	0.25	1.31
	28			0.26	0.72			0.25	1.23
	4	0.31	0.03				0.80		1.14
	29			0.26		0.04	0.80		1.10
	18		0.03	0.26			0.80		1.09
	35					0.04	0.80	0.25	1.09
	25		0.03				0.80	0.25	1.08
	10	0.31			0.72	0.04			1.07
	2	0.31	0.03		0.72				1.06
	26			0.26	0.72	0.04			1.02
	16		0.03	0.26	0.72				1.01
	22		0.03		0.72			0.25	1.00
	23		0.03			0.04	0.80		0.87
	9	0.31		0.26				0.25	0.82
	20		0.03		0.72	0.04			0.79
	14	0.31				0.04		0.25	0.60
	7	0.31		0.26		0.04			0.61
	1	0.31	0.03	0.26					0.60
	5	0.31	0.03					0.25	0.59
	30			0.26		0.04		0.25	0.55
	19		0.03	0.26				0.25	0.54
	3	0.31	0.03			0.04			0.38
	17		0.03	0.26		0.04			0.33
	24		0.03			0.04		0.25	0.32

Table 6. Relative importance.

	Economic	Legal	Political	Social	Cultural	Ethical	Environmental
Appearing	1.55	0.00	0.52	5.04	0.12	4.80	1.00
Not appearing	3.10	0.45	3.38	5.76	0.48	7.20	2.75

This leads to two complementary positions: on the one hand, to the induction of the current concept of sustainability as that resulting from the two most commonly used approaches within the political-social-ethical dimension and to complement it with the definition of greatest weighting, as shown in Table 7.

Table 7. Induction.

Holling (2001)	Process of creating, innovating and experimenting in order to increase the adaptive capacities of ecosystems and of individuals and human societies as a whole, in a changing context of disturbance, uncertainty and surprise	4557	Political Social Ethical
Montensen (2000, p. 29)	Represents a new paradigm based on a model of citizenship that must be and needs to be shaped and informed	271	Political Social Ethical

4. Conclusions

It is clear from the study that the dimensions that currently contribute most to the dynamics of sustainability processes are the political, social and ethical. Indeed, each of these dimensions makes a significant individual contribution (Table 2), and jointly their weighting fraction is 44.6% (Table 4). According to the distribution set out in Table 3, three of the original seven categories are sufficient to establish the definition of sustainability due to their large number of UAs and weighting in the Google count.

The qualitative study carried out in this work makes it possible to conclude that the trend of current sustainability processes is to move toward environmental contributions—albeit in the very long term due to existing weighting differences—while the legal and cultural dimensions are excluded, being implicitly residual in character.

Sustainability describes a phenomenon imbued with political interests, despite the lack of public policies that ensure its deployment and normal development, which should rather emphasize strong social commitment and defense of the principles that guide human behavior. In other words, instead of being a common and widespread commitment, it is evident that sustainability processes are used as subordinate elements of economic growth and are complicit in substantial environmental degradation. This is shown by the fact that, contrary to the a priori societal expectation, the environmental dimension is not included among the most important definitions, thereby conditioning the implementation, deployment and environmental protections offered by sustainability processes. It should be noted that if sustainability processes are not managed in a responsible manner, there is a danger of generating new adverse environmental and social impacts that may arise from the interrelationships between all dimensions that are not reflected in the concepts currently used and analyzed by experts.

At some point, the question will have to be asked how we, as active citizens, might address, critically and consciously, the social, economic and political order that determines our societies and our lifestyles.

Author Contributions: Conceptualization, A.L.G., J.Á.C.M., A.C.V.-T. and F.R.; methodology, A.L.G., J.Á.C.M., A.C.V.-T. and F.R.; software, A.L.G., J.Á.C.M., A.C.V.-T. and F.R.; validation, A.L.G., J.Á.C.M., A.C.V.-T. and F.R.; formal analysis, A.L.G., J.Á.C.M., A.C.V.-T. and F.R.; investigation, A.L.G., J.Á.C.M., A.C.V.-T. and F.R.; resources, A.L.G.; data curation, A.L.G. and J.Á.C.M.; writing—original draft preparation, A.L.G., J.Á.C.M., A.C.V.-T. and F.R.; writing—review and editing, A.L.G., J.Á.C.M., A.C.V.-T. and F.R.; visualization, A.L.G., J.Á.C.M., A.C.V.-T. and F.R.; supervision, A.L.G.; project administration, A.L.G.; funding acquisition, A.L.G. and J.Á.C.M. All authors have read and agreed to the published version of the manuscript.

Funding: Research project UTM 2021-23-45.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

The table set out below shows the source of the definitions of the processes of sustainability, the concepts related to each author and the dimension to which each belongs.

Table A1. Categorization of the concepts of sustainability.

Source	Concept	Frequency	Dimension
Acosta, Lovato and Buñay (2018) [32]	Process improvement and optimization of resources that influence cost reduction	7	Economic
Agyeman (2005) [33]	The need to ensure a better quality of life for all, now and in the future, in a fair and equitable way, while living within the limits of ecosystem capacity	774	Economic Political Social Environmental Ethical
Allende, (1995) [34]	Reflecting a policy and strategy of continuous economic and social development that is not detrimental to the environment or natural resources on whose quality the continuity of human activity and development depends	74	Economic Political Social Environmental Cultural Ethical
Alonso-Almeida, Marimon and Llach (2015, p. 140) [35]	Meeting the needs of today's generation without sacrificing the ability of future generations to meet their own needs	61	Social Ethical
Aragonés, Izurieta and Raposo (2003) [36]	Discourse involving political, social and economic actors	49	Economic Political Social
Artaraz (2002) [37]	Productive systems that use only renewable resources and energies, and do not produce waste, since they return to nature (through compost) or become an input for another manufactured product	342	Economic Environmental Ethical
Austermühle (2012) [38]	Having no negative impact on the global environment, society or economy through environmentally friendly activities, ensuring that all processes, products and operations consider the environmental challenges while producing an economic benefit	39	Economic Social Environmental
Ávila (2018) [39]	Exploiting natural resources in a way that remains within the limits of regeneration and natural growth by planning the exploitation of resources and specifying the effects that exploitation will have on the whole ecosystem	92	Economic Environmental Ethical
Ayres (1996) [40]	Argues that the concept of sustainability guides how humanity must act in relation to nature and be responsible for its own generations	100	Environmental Ethical
Azqueta (1994) [41]	Subordination of nature conservation to economic growth	323	Economic Environmental
Banco Mundial (1999) [42]	The ability of a project to maintain an acceptable level of flow of profits throughout its economic life		Economic Legal
Barcellos (2010) [43]	It is an ongoing process that depends on the commitment to pursuing the goals set and having strategies that generate a long-term competitive advantage	1	Economic
Bermejo (2008) [44]	Ensuring that the needs of the present are met without compromising the ability of future generations to meet their own	1	Social Ethical
BNP Paribas (2009) [45]	Point at which the economic, social and environmental demands of a business can be balanced with those of the society in which it develops, thereby caring for the needs of the future	3	Economic Social Environmental Ethical

Table A1. *Cont.*

Source	Concept	Frequency	Dimension
Boada, Rocchi and Kuhndt (2005) [46]	It refers not only to environmental aspects, but also involves social responsibility and economic viability, which reflects the complexity of the concept, since restricting it to the environmental gives an erroneous as well as a partial or incomplete perspective	6	Economic Social Environmental Ethical
Buarque (1994) [47]	Qualitative and quantitative process of social change that balances environmental conservation and social equity over time and space	1	Economic Social Environmental
Bybee (1991, p. 151) [48]	It is the outstanding central unifying idea at this time in human history and requires the consideration of all interconnected problems on a global scale	383	Social Ethical
Camacho-Ruiz, Carrillo-Reyes, Rioja-Paradela, and Espinoza-Medinilla (2016) [49]	Sustaining development while maintaining its physical and vital sustenance	41	Economic Ethical
Carpintero (1999) [50]	The lasting relationship of every socioeconomic system with its ecosystem	2	Economic Social
Carrizosa (1998) [51]	Inherent property of a process that perpetuates its existence within a given system	31	Ethical
Carvalho (1994) [52]	Process of transformation in which the exploitation of resources, the direction of investments, the orientation of technological development and institutional change harmonize and reinforce the present and future potential, with the aim of meeting human needs and aspirations	12	Economic Political Social
Cepal and Pnuma (2002) [53]	Recognition of the role of the environment and natural resources in ensuring economic progress	15	Economic Environmental
CNUMAD (1992) [54]	Integration of economic, social, cultural, political and ecological factors	5	Economic Political Social Environmental Cultural
Colucci-Gray, Camino, Barbiero, and Gray (2006, p. 244) [55]	A system of values that has emerged in recent years as a result of a new sensitivity toward the earth and related to "ethics of the planet"	261	Environmental Ethical
ConteGrand and Deliar (2018, p. 66) [56]	Maximum social well-being, defined based on the values of society	17	Social Cultural Ethical
Costanza (1997) [57]	Viability of the complex interaction between two dynamic systems, the socioeconomic and the ecosystem	59	Economic Social
Contreras and Rojas (2015, p. 77) [58]	Long-term value creation	51	Economic
Conway (1993, p. 380) [59]	Ability of a system to maintain its productivity despite disturbances	220	Economic Political
Cutter-Mackenzie and Smith (2003) [60]	It is not only about understanding problems, but making sense of everyday experiences and doing so both in the personal and social spheres and in the global framework	261	Social Ethical
Daly (1990) [61]	Viability of a socioeconomic system over time	2182	Economic Social
Díaz and Camejo (2015) [62]	Viable development over time whose essential condition is that the capabilities of the socioeconomic system not increase and may be available for the generations to come	217	Economic Social Ethical

Table A1. *Cont.*

Source	Concept	Frequency	Dimension
Ebel and Kissmann (2011, p. 72) [63]	The needs of future and current generations		Social Ethical
Edwars (2005) [64]	Diverse, multicultural, multiperspective and global revolution, built around four dimensions, ecological, economic, fairness and education	14	Economic Cultural Ethical
Ehrenfeld (2005) [65]	Possibility of allowing all forms of life to always thrive		Ethical
Elizalde (2004) [66]	It is a collective task with no pre-established conditions since it is a conceptual construction that requires debate and the participation of all the actors involved. To achieve this will demand a huge political and cultural effort of all humankind	238	Political Social Cultural Ethical
Estévez (2013) [67]	Addressing current needs without compromising the ability of future generations to meet their own, ensuring the balance between economic growth, environmental care and social well-being	70	Economic Social Environmental Ethical
Feil and Schreiber (2017, p. 673) [68]	Having the worry about future existence of natural resources to support the continuation of human life as its foundation	33	Social Environmental Ethical
Fernández and Gutiérrez (2013, p. 122) [69]	Improving the quality of human life by living within the load-bearing or sustaining capacity of supporting ecosystems	117	Social Environmental Ethical
Gallopín (1996, p. 105) [70]	Set of basic system attributes viewed as a whole	323	Ethical
Gallopín (2003) [9]	Attribute of systems open to interactions with the external world. It is not a fixed state of constancy, but the dynamic preservation of the essential identity of the system amid permanent changes	713	Cultural Ethical
Garcés (2000) [71]	It is a new way of understanding the relationships between people and their environment from a social perspective that involves prioritizing the psychological and social satisfaction of the individual and their quality of life over algorithmic processes	5	Political Social Ethical
Gómez (2015) [72]	It involves comparing various scenarios involving resource consumption, the most sustainable being that which can be made effective with the greatest efficiency and at the lowest possible cost in resources	1	Economic
González, Montes, Santos, and Monedero (2008) [73]	Basic principle or premise of global sustainable development, in which a society's primary objective is to meet its basic needs in such a way that it can maintain, now and in the future, acceptable levels of social well-being	9	Social Ethical
Green Facts (2021) [74]	A characteristic or state whereby the needs of the present and local population can be met without compromising the ability of future generations or populations in other locations to meet their needs	1	Economic Social Environmental Ethical
Guzmán and Alonso (2007) [75]	Exploitation of natural resources that does not put in jeopardy their availability to future generations and that leads to a sense of nature as a resource of the liberal economy	127	Economic Social Environmental
González and De Lázaro y Torres (2005) [76]	It is the quest for urban development that does not degrade the environment and provides quality of life for citizens	4	Economic Political Social Environmental Ethical
Hart and Milstein (2003) [77]	Creating value at the level of strategies and practices	2115	Ethical

Table A1. *Cont.*

Source	Concept	Frequency	Dimension
Hicks (1946, p. 27) [78]	Maximum amount of resources that a person can consume in a period and still be as well off at the end of the period as s/he was at the beginning	72	Economic Legal Political Social Environmental Cultural Ethical
Holling (2001) [79]	Process of creating, innovating and experimenting in order to increase the adaptive capacities of ecosystems and individuals and human societies as a whole in a changing context of disturbance, uncertainty and surprise	4557	Political Social Ethical
Idrovo-Carlier, S. and Torres-Castillo (2017) [80]	It refers to the individual. Human sustainability is achieved when people are able to harmonize all areas in which they develop to reach their personal fulfillment without sacrificing their quality of life	1	Social Ethical
Jimenez (1998, p. 61) [81]	Principle of complementarity and incompatibility between growth and nature	4	Economic Environmental Ethical
Kammerbauer (2001) [82]	A policy objective that falls within the ethics of responsibility	49	Political Social Ethical
Kates et al. (2001) [83]	Is caring for the earth, its living biota and its people	3504	Social Environmental Ethical
Khandker, Baqui, and Zahed (1995) [84]	The program's ability to carry out its activities and services on an ongoing basis, in the quest for the fulfillment of its objectives	556	Ethical
Lalangui, Espinoza, and Pérez (2017, p. 148) [85]	An integral part of social responsibility	58	Social Ethical
Langenwalter (2009, p. 10) [86]	Sustainability involves complex and changing environmental dynamics that affect human livelihoods and well-being, with intersecting ecological, economic and sociopolitical dimensions, both globally and locally	15	Economic Legal Social Environmental
Linares (2013) [87]	Creating value on aggregate. Individuals are sustainable when they add more value than they subtract. The same is true of an enterprise, a country or a specific activity	21	Economic Ethical
López (2006) [88]	Intergenerational protection of natural resources, so that future generations can enjoy them in the same conditions and at the same level as previous generations	2	Social Environmental
López, López and Ancona (2005) [89]	Process that harmonizes economic growth, the preservation of natural resources, the reduction of environmental deterioration, social equity, all in a political context operating at all levels: local, regional, national and global	100	Economic Legal Political Social Environmental Cultural Ethical
Lufiego and Rabadán (2000) [11]	Generic concept that is defined as the viability of a socioeconomic system over time	69	Economic Social Ethical
Macedo (2005) [90]	It is the development that meets the needs of the present generation without compromising the ability of future generations to meet their own needs	113	Social Ethical
Marcén (2009) [91]	Concept associated with the realization that the world is not as broad and limitless as we had believed	2	Ethical

Table A1. *Cont.*

Source	Concept	Frequency	Dimension
Meadows, Meadows, and Randers (1991) [92]	Persisting over generations and able to look to the future with enough flexibility and wisdom as not to undermine physical or social systems of support	3937	Social Ethical
Millán, Hidalgo, and Arjona (2015) [93]	Responsibility related to the environment, natural resources and Indigenous communities in various rural areas	58	Social Environmental Cultural
Mitcham (1995) [94]	Exchanging unsustainable trends and intrinsic contradictions for the interaction between economic development and environmental deterioration	242	Economic Environmental Ethical
Mokate (2001) [95]	The ability to be maintained or sustained. An initiative is sustainable if it persists over time	391	Ethical
Mokate (2004) [96]	The ability of an initiative to maintain an adequate flow of resources to ensure the continuity of expected effects	3	Ethical
Montensen (2000, p. 29) [97]	Represents a new paradigm based on a model of citizenship that must be and needs to be shaped and informed	271	Political Social Ethical
Naredo (1996, p. 133) [98]	Viability of the relationship between a socioeconomic system and an ecosystem	333	Economic Social Environmental
Nogales (2006) [99]	Having as a point of reference the natural order and the use of the knowledge we have acquired about natural and ecological systems	43	Social Environmental Ethical
O'Connor (1994, p. 411) [100]	Relationship between a socioeconomic system and an ecosystem whose rate of entropy is compatible with maintaining that relationship over time	14	Economic Social Ethical
O'Connor (2006) [101]	Focuses on interfaces, interactions and interdependencies between the economic, social and environmental spheres, with the mediation of the political dimension	220	Economic Political Social Environmental
Organización de las Naciones Unidas (2018) [102]	Processes of inclusion and stability	13	Economic Ethical
O'Riordan (1988) [103]	It is a phenomenon that encompasses ethical standards related to the survival of all living things, the rights of future generations and the institutions responsible for ensuring that such rights are taken into account in policies and actions	469	Legal Political Social Ethical
Partridge (2011) [104]	Has many definitions, with the three most common being an activity that can be continued indefinitely without causing harm; doing unto others as you would have them do unto you; and meeting a current generation's needs without compromising those of future generations	47	Ethical
Pearce (1997, p. 11) [105]	It requires that the conditions necessary for equal access to the resource base be met for the benefit of each generation	670	Social Ethical
Pearce, Albritton, Grant, Steed, and Zelenika (2012, p. 44) [106]	Appropriate technology, social entrepreneurship, service learning and international development to focus on what they do best	58	Economic Social
Pérez (2012, p. 142) [107]	Paradigm that seeks to reconcile the conflict between the production of goods and services for society (development) and the environmental supply available in terms of natural resources and ecosystem services	151	Economic Political Social Ethical
Pezzey (1992) [108]	Nondeclining utility of a representative member of society for millennia into the future.	443	Social Ethical

Table A1. *Cont.*

Source	Concept	Frequency	Dimension
Pezzey and Toman (2002, p. 14) [109]	Concern for intergenerational equity in the long-term decision making of a society.	186	Social Ethical
Pinillos and Fernández (2011) [110]	Creating long-term value by leveraging opportunities and effectively exploiting the risks inherent in economic, environmental and social development	146	Economic Social Environmental
Plasencia, Marrero, Bajo, and Nicado (2018, p. 69) [111]	Development that meets the needs of these generations without compromising the ability of future generations to meet their needs	30	Economic Social Ethical
Quiroga (2001) [112]	Leaving to future generations a stock of natural capital equal to that available to present generations, that is, sustainability is seen as a process of maintaining natural capital	332	Economic Social Ethical
Rees (1990, p. 22) [113]	Opportunity for a return to community values, local control over resources, community-based development and other forms of decentralized government	329	Legal Political Social Ethical
Robilliard (2006) [114]	That which, in its development, takes into account economic, social and environmental aspects with a democratic attitude that involves the participation of those affected by the process	22	Economic Social Environmental
Rodríguez-Ariza, Frías, and García (2014, p. 6) [115]	Economic, social and environmental impact of business activities over a given period of time	69	Economic Social Environmental
Rodríguez and Govea (2006) [116]	Anthropocentric and comprehensive concept that provides well-being to present and future generations without environmental-social-economic deterioration	37	Economic Social Environmental Ethical
Rojas (2009) [117]	Sustainability can be understood as the balance between the environmental, economic and social that translates into urban quality of life	8	Economic Social Environmental Ethical
Rueda (2002) [118]	Process of social change wherein the exploitation of resources, the orientation toward technological development and institutional reforms are carried out harmoniously, expanding the current and future potential for the satisfaction of human needs and aspirations	8	Economic Social Ethical
Schreiner, Henriksen, and Hansen (2005, p. 13) [119]	Moral education on the responsibility for social action with a perspective on both today and tomorrow	173	Social Ethical
Toro (2007) [6]	Maintenance of a phenomenon or dynamic process over time, within margins that condition its viability. This idea of its ability to last is related to its withstanding possible alterations caused by external elements and/or sudden internal changes	31	Ethical
UICN, PNUMA, WWF (1991) [120]	Maintaining the load-bearing capacity of an ecosystem over the course of the relationship between society and the ecosystem	17	Social Environmental Ethical
Uribe, Vargas, and Merchán (2018, p. 61) [121]	Permanence of results overtime	9	Economic
Vasconez and Torres (2018, p. 49) [122]	Anthropological and technocentric concept, functioning in the development of productive forces and indefinite growth, which puts a price on nature	11	Economic Environmental
Vaitheeswaran (2002) [123]	It is what we owe future generations and how we can reconcile that with what we owe the poorest among us today	46	Economic Social Ethical

Table A1. *Cont.*

Source	Concept	Frequency	Dimension
Vilches and Gil (2013) [124]	It requires holistic approaches that take into account all the problems that characterize the planetary emergency because they are closely interconnected and mutually empowered	59	Social Environmental Ethical
Vilches and Gil (2015, p. 39) [125]	Addressing the current global emergency situation and driving societies' transition	69	Social Ethical
Willis (2003) [126]	Communication tool, useful for dialogue between stakeholders	1746	Economic
World Bank (2003) [127]	Improve human well-being over time. The poor and the vulnerable must have much greater access to assets for growth to be sustainable and for the world to avoid social unrest	20	Economic Social Ethical
World Commission on Environment and Development (1987) [128]	Meeting the needs of the present without compromising the ability of future generations to meet their own	503	Economic Ethical

References

1. Sánchez Fernández, G. Análisis De La Sostenibilidad Agraria Mediante Indicadores Sintéticos: Aplicación Empírica Para Sistemas Agrarios. Ph.D. Thesis, Universidad Politécnica de Madrid, Madrid, Spain, 16 July 2009.
2. Huisingh, D. New challenges in education for sustainable development. *Clean Technol. Environ. Policy* **2006**, *8*, 3–8. [CrossRef]
3. Ahmed, Z.; Alam, R.; Akter, S.A.; Kadir, A. Environmental sustainability assessment due to stone quarrying and crushing activities in Jaong, Sylhet. *Environ. Monit. Assess.* **2020**, *192*, 778. [CrossRef] [PubMed]
4. Cuspilici, A.; Monforte, P.; Ragusa, M.A. Study of Saharan dust inuence on PM10 measures in Sicily from 2013 to 2015. *Ecol. Indic.* **2017**, *76*, 297–303. [CrossRef]
5. Chen, W.K.; Nalluri, V.; Lin, M.L.; Lin, C.T. Identifying Decisive Socio-Political Sustainability Barriers in the Supply Chain of Banking Sector in India: Causality Analysis Using ISM and MICMAC. *Mathematics* **2021**, *9*, 240. [CrossRef]
6. Toro Sánchez, F.J. El desarrollo sostenible: Un concepto de interés para la geografía. *Cuadernos Geográficos* **2007**, *40*, 49–181. [CrossRef]
7. Ríos-Osorio, L.; Cruz-Barreiro, I.; Welsh-Rodriguez, C. The Concept of Sustainable Development from an Ecosystem Perspective: History, Evolution, and Epistemology. In *Ecological Dimensions for Sustainable Socio Economic Development. Southampton*; Yañez-Arancibia, A., Dávalos-Sotelo, R., Day, J., Reyes, E., Eds.; WIT Press: Boston, MA, USA, 2013; Chapter 2; pp. 29–46.
8. Brundtland, G.H. Informe de la Comisión Mundial Sobre Medio Ambiente y el Desarrollo: Nuestro futuro Común. Documentos de las Naciones Unidas, 416. 1987. Available online: https://bit.ly/3w1GHKG (accessed on 8 February 2021).
9. Gallopín, G. *Sostenibilidad y Desarrollo Sostenible: Un Enfoque Sistémico*; CEPAL, Serie Medio Ambiente y Desarrollo: Santiago, Chile, 2003; pp. 1–44.
10. Salas-Zapata, W.A.; Ríos-Osorio, L.A.; Del Castillo, J.A. Bases conceptuales para una clasificación de los sistemas socioecológicos de la investigación en sostenibilidad. *Rev. Lasallista Investig.* **2011**, *8*, 136–142.
11. Lufiego García, M.; Rabadán Verfgara, J.M. La evolución del concepto de sostenibilidad y su introducción en la enseñanza. *Revista de Investigación y Experiencias Didácticas* **2000**, *18*, 473–486.
12. Luque, A.; Casado, F. Public Strategy and Eco-Social Engagement in Latin American States: An Analysis of Complex Networks Arising from Their Constitutions. *Sustainability* **2020**, *12*, 8558. [CrossRef]
13. Luque, A.; Herrero-García, N. How corporate social (ir)responsibility in the textile sector is defined, and its impact on ethical sustainability: An analysis of 133 concepts. *Corp Soc. Responsib. Environ. Manag.* **2019**, *26*, 1285–1306. [CrossRef]
14. National Environmental Policy Act. EPA United States Environmental Protection Agency. Available online: https://www.epa.gov/nepa (accessed on 2 February 2021).
15. United Nations. Perder o Desperdiciar Menos Comida Ayudará a Conservar el Medio Ambiente. Available online: https://news.un.org/es/story/2019/10/1463751 (accessed on 14 December 2020).
16. ZWIA—Zero Waste International Alliance. Working towards a World without Waste. Available online: http://zwia.org/ (accessed on 2 February 2021).
17. Estado Plurinacional de Bolivia. Asamblea Legislativa Plurinacional. Ley 300. Ley Marco de la Madre Tierra y Desarrollo Integral para Vivir Bien. Available online: http://www.fao.org/fileadmin/user_upload/FAO-countries/Bolivia/docs/Ley_300.pdf (accessed on 4 February 2021).
18. Luque, A. Promoción del hiperconsumo textil transnacional: La moda y el exceso como leitmotiv. *Chasqui. Rev. Latinoam. Comun.* **2017**, *134*, 83–104. [CrossRef]
19. Dominish, E.; Florin, N.; Teske, S. Responsible Minerals Sourcing for Renewable Energy. In *Report Prepared for Earthworks by the Institute for Sustainable Futures*; University of Technology Sydney: Sydney, Australia, 2019.
20. King, S.; Boxall, N.J.; Bhatt, A.I. *Australian Status and Opportunities for Lithium Battery Recycling*; CSIRO: Canberra, Australia, 2018.

21. Visser, W. Revisiting Carroll's CSR pyramid: An African Perspective. In *Corporatecitizenship in Developing Countries*; Pedersen, E.R., Huniche, M., Eds.; Copenhagen Business School Press: Copenhagen, Danmark, 2006; pp. 29–56.
22. Torugsa, N.; O'Donohue, W.; Hecker, R. Proactive CSR: An empirical analysis of the role of its economic, social and environmental dimensions on the association between capabilities and performance. *J. Bus. Ethics* **2013**, *115*, 383–402. [CrossRef]
23. Ley de Economía Social/Gobierno de España. Ley 5/2011, de 29 de Marzo, de Economía Social. Available online: https://www.boe.es/buscar/doc.php?id=BOE-A-2011--5708 (accessed on 10 February 2021).
24. Rostow, E. *Planteamiento Para la Libertad*; Editorial Legal Omeba: Buenos Aires, Argentina, 1962.
25. Luque, A. Corruption in the transnational textile industry: An exception or the rule? *Empresa y Humanismo* **2018**, *21*, 123–184. [CrossRef]
26. Leguizamón, A.W. *Derecho Económico*; Ediciones Doctrina y Ley: Bogotá, Colombia, 2002.
27. Luque, A.; de Pablos, C. Factors that promote the lack of ethics in the production and logistics practices of transnational companies. *Lan Harremanak* **2016**, *34*, 336–370. [CrossRef]
28. Luque, A.; Guamán, A. Transnational textile outsourcing: Exceptional or standard? *J. Bus. Glob.* **2021**, *27*, 143–170. [CrossRef]
29. UNESCO. Conferencia Mundial Sobre las Politicals Culturales. Available online: http://unesdoc.unesco.org/images/0005/000525/052505sb.pdf02/02/1982 (accessed on 2 February 2021).
30. United Nations Environment Programme. *Making Peace with Nature. A Scientific Blueprint to Tackle the Climate, Biodiversity and Pollution Emergencies*; United Nations Environment Programme: Nairobi, Kenya, 2021. Available online: https://bit.ly/3uVieVR (accessed on 8 February 2021).
31. Luque, A.; Jiménez-Sánchez, A. Textile sustainability processes: Success, or a new way to contaminate under a friendly paradigm. *J. Incl.* **2018**, *6*, 259–288.
32. Acosta Véliz, M.; Lovato Torres, S.; Buñay Cantos, J. La responsabilidad social corporativa y su rol en las empresas ecuatorianas. *Lasallista Investig.* **2018**, *15*, 105–117. [CrossRef]
33. Agyeman, J. *Environmental Justice and Sustainable Communities*; New York University Press: New York, NY, USA, 2005.
34. Allende Landa, J. Desarrollo sostenible. De lo global a lo local. Ciudad y Territorio. *Estud. Territ.* **1995**, *III*, 267–281.
35. Alonso-Almeida, M.; Marimon, F.; Llach, L. Difusión de las memorias de sostenibilidad en Latinoamérica: Análisis territorial y sectorial. *Estud. Gerenc.* **2015**, *31*, 139–149. [CrossRef]
36. Aragonés, J.; Izurieta, C.; Raposo, G. Revisando el concepto de desarrollo sostenible en el discurso social. *Psicothema* **2003**, *15*, 221–226.
37. Artaraz, M. Teoría de las tres dimensiones de desarrollo sostenible. *Rev. Ecosistemas* **2002**, *11*, 1–6. [CrossRef]
38. Austermühle, S. *Sostenibilidad y Ecoeficiencia en la Empresa Moderna*; Universidad Peruana de Ciencias Aplicadas: Lima, Peru, 2012.
39. Ávila, P. La sustentabilidad o sostenibilidad: Un concepto poderoso para la humanidad. *Tabula Rasa* **2018**, *28*, 409–423. [CrossRef]
40. Ayres, R. Statistical measures of unsustainability. *Ecol. Econ.* **1996**, *16*, 239–255. [CrossRef]
41. Azqueta, D. *Valoración Económica de la Calidad Ambiental*; McGraw-Hill e Interamericana de España: Madrid, Spain, 1994.
42. Banco Mundial. *Sustainability of Projects: Review of Experience in the Fertilizer Subsector*; World Bank Group: Washington, DC, USA. Available online: https://bit.ly/3g6JhrK (accessed on 15 January 2021).
43. Barcellos, L. *Modelos de Gestión Aplicados a la Sostenibilidad Empresarial*; Universidad de Barcelona: Barcelona, Spain, 2010.
44. Bermejo Gómez de Segura, R. *Del Desarrollo Sostenible Según Brundtland a la Sostenibilidad Como Biomimesis*; Universidad del País Vasco: Bilbao, Spain, 2008.
45. BNP Paribas Real State. *El Glosario de la Sostenibilidad*; MediaResponsable, S.L: Barcelona, Spain, 2009.
46. Boada Ortiz, A.; Rocchi, S.; Kuhndt, M. *Negocios y Sostenibilidad más Allá de la Gestión Ambiental*; Politécnico Grancolombiano: Bogota, Colombia, 2005.
47. Buarque, S. *Desenvolvimento Sustentável da Zona da Mata de Pernambuco*; IICA/SEPLAN: Recife, Brazil, 1994.
48. Bybee, R. Planet earth in crisis. *Am. Biol. Teach.* **1991**, *53*, 146–153. [CrossRef]
49. Camacho-Ruiz, E.; Carrillo-Reyes, A.; Rioja-Paradela, T.M.; Espinoza-Medinilla, E.E. Indicadores de sostenibilidad para el ecoturismo en México: Estado actual. *LiminaR* **2016**, *XIV*, 156–168. [CrossRef]
50. Carpintero, O. *Entre la Economía y la Naturaleza*; Fundación 1º de Mayo: Madrid, Spain, 1999.
51. Carrizosa, J. Construcción de la Teoría de la Sostenibilidad. *Revista Transición Convivencia y Sostenibilidad* **1998**, *28*, 27–32.
52. Carvalho, H. *A participaçao e Organizaçao Consentida Como uma das Dimensões da Cidadania*; Informe Integral; IICA/SEPLAN: Brasilia, Brazil, 1994.
53. Cepal y Pnuma. *La Sostenibilidad del Desarrollo en América Latina y el Caribe. Desafíos y Oportunidades*; Libros de Cepal: Santiago, Chile, 2002.
54. CNUMAD. Programa 21. Conferencia de las Naciones Unidas sobre el Medio Ambiente y el Desarrollo. Rio de Janeiro. Available online: https://www.un.org/spanish/esa/sustdev/agenda21/riodeclaration.htm (accessed on 18 January 2021).
55. Colucci-Gray, L.; Camino, E.; Barbiero, G.; Gray, D. From Scientific Literacy to Sustainability: An Ecological Framework for Education. *Sci. Educ.* **2006**, *90*, 227–252. [CrossRef]
56. ConteGrand, M.; Deliar, V. Desarrollo sostenible y conceptos verdes. *Problemas del Desarrollo* **2018**, *192*, 63–75. [CrossRef]
57. Constanza, R. La economía Ecológica de la Sostenibilidad: Invertir en Capital Natural. In *Medio Ambiente y Desarrollo Sostenible: Más Allá del Informe Brundtland*; Goodland, R., Ed.; Trota: Madrid, Spain, 1997; pp. 103–114.

58. Contreras, O.; Rojas, I. Teletrabajo y sostenibilidad empresarial. Una reflexión desde la gerencia del talento humano en Colombia. *Suma de Negocios* **2015**, *6*, 74–83. [CrossRef]
59. Conway, G. Sustainable Development: Differing Perspectives of Ecologists and Economists. *World Dev.* **1993**, *16*, 373–384. [CrossRef]
60. Cutter-Mackenzie, A.; Smith, R. The missing paradigm in environmental education. *Environ. Educ. Res.* **2003**, *9*, 497–524. [CrossRef]
61. Daly, H. Toward some operational principles of sustainable development. *Ecol. Econ.* **1990**, *2*, 1–6. [CrossRef]
62. Díaz Geniz, A.; Camejo, M. *Epistemología y Educación: Articulaciones y Convergencias*; Espacio Interdisciplinario: Montevideo, Uruguay, 2015.
63. Ebel, R.; Kissmann, S. Desarrollo sostenible: La investigación en un contexto intercultural. *Ra Ximhai* **2011**, *7*, 69–79. [CrossRef]
64. Edwars, A. *The Sustainability Revolution*; New Society Publishers: Gabriola, BC, Canada, 2005.
65. Ehrenfeld, J. The roots of sustainability. *MIT Sloan Manag. Rev.* **2005**, *46*, 23–26.
66. Elizalde, A. Paradojas y punto de fuga para una sostenibilidad posible. *Polis* **2004**, *9*, 1–20.
67. Estévez, R. Las 10 Definiciones Ecointeligentes que Debes Conocer. Available online: https://www.ecointeligencia.com/2013/02/10-definiciones-sostenibilidad/ (accessed on 10 December 2020).
68. Feil, A.; Schreiber, D. Sustainability and sustainable development: Unraveling overlays and scope of their meanings. *Cad. EBAPE. BR* **2017**, *14*, 667–681. [CrossRef]
69. Fernández, L.; Gutiérrez, M. Bienestar social, económico y ambiental para las presentes y futuras generaciones. *Información Tecnológica* **2013**, *24*, 121–130. [CrossRef]
70. Gallopín, G. Branching Futures and Energy Projections. *Renew. Energy Dev.* **1996**, *10*, 101–107.
71. Garcés, J. *La Nueva Sostenibilidad Social*; Ariel: Barcelona, Spain, 2000.
72. Gómez Jiménez, M.L. *Sostenibilidad Social y Rehabilitación de Vivienda: Especial Referencia a Andalucía*; Editorial Dykinson S.L.: Madrid, Spain, 2015.
73. González, J.; Montes, C.; Santos, I.; Monedero, C. Invirtiendo en capital natural: Un marco para integrar la sostenibilidad ambiental en las Politicals de cooperación. *Rev. Ecosistemas* **2008**, *17*, 52–69.
74. Green Facts Millennium Ecosystem Assessment. Available online: https://www.greenfacts.org/glossary/pqrs/sustainability.htm (accessed on 13 January 2021).
75. Guzmán, G.I.; Alonso, A.M. La investigación participativa en agroecología: Una herramienta para el desarrollo sustentable. *Ecosistemas* **2007**, *16*, 24–36.
76. González Gonzalez, M.J.; de Lázaro y Torres, M.L. La localización de la población inmigrante en España a través del Censo de la Población del 2001. Principales características. *Cuadernos Geográficos* **2005**, *36*, 35–49. [CrossRef]
77. Hart, S.; Milstein, M. Creating sustainable value. *Acad. Manag. Exec.* **2003**, *17*, 56–69. [CrossRef]
78. Hicks, J. *Value and Capital*; Clarendon Paperbacks: London, UK, 1946.
79. Holling, C.S. Understanding the complexity of economic, ecological, and social systems. *Ecosystems* **2001**, *4*, 390–405. [CrossRef]
80. Idrovo-Carlier, S.; Torres-Castillo, C. *El Lado Humano de la Sostenibilidad. Reflexiones Desde lo Privado y lo Público*; Universidad de la Sabana, Chía: Cundinamarca, Colombia, 2017.
81. Jimenez Herrero, L.M. Indicadores de desarrollo. Tras la fórmula de la sostenibilidad. *Ecosistemas* **1998**, *24*, 58–63.
82. Kammerbauer, J. Las dimensiones de la sostenibilidad: Fundamentos ecológicos, modelos paradigmáticos y senderos. *Interciencia* **2001**, *26*, 353–359.
83. Kates, R.W.; Clark, W.C.; Corell, R.; Hall, J.M.; Jaeger, C.C.; Lowe, I.; Mccarthy, J.J.; Schellnhuber, H.J.; Bolin, B.; Dickson, N.M.; et al. Sustainability Science. *Science* **2001**, *292*, 641–642. [CrossRef]
84. Khandker, S.; Baqui, K.; Zahed, K. *Performance and Sustainability*; Documento de Discusión No. 306; Banco Mundial: Washington, DC, USA, 1995.
85. Lalangui, J.; Espinoza Carrión, C.R.; Pérez Espinosa, M.J. Turismo sostenible, un aporte a la responsabilidad social empresarial: Sus inicios, características y desarrollo. *Universidad y Sociedad* **2017**, *9*, 148–153.
86. Langenwalter, G.A. Planet first (incorporating sustainability in corporate strategy). *Strateg. Dir.* **2010**, *26*. [CrossRef]
87. Linares, P. El concepto marco de sostenibilidad: Variables de un futuro sostenible. In *¿Es sostenible el mundo en que vivimos?: Un enfoque interdisciplinar*; Bedate, C., Ed.; Universidad Pontificia Comillas: Madrid, Spain, 2013; pp. 11–32.
88. López Astudillo, A. La ruta de la sostenibilidad. *Estud. Gerenc.* **1998**, *69*, 37–42.
89. López Ricalde, C.D.; López Hernández, E.S.; Ancona Peniche, I. Desarrollo sustentable o sostenible: Una definición conceptual. *Horiz. Sanit.* **2005**, *4*. [CrossRef]
90. Macedo, B. *El Concepto de Sostenibilidad*; Oficina Regional de Educación para América Latina y el Caribe UNESCO: Santiago, Chile, 2005.
91. Marcén Albero, C. El Gran Desafío: Diseñar el Papel de la Escuela en la Apuesta Colectiva por la Buena Gobernabilidad del Agua. In *Agua y Sostenibilidad. Recursos, Riesgos y Remedios*; Moreno González, A., López Martínez, C., Eds.; Ministerio de Educación Cultura y Deporte: Madrid, Spain, 2009; pp. 231–250.
92. Meadows, D.; Meadows, D.; Randers, J. *Mas Allá de Los Límites del Crecimiento*; Editorial El País Aguilar: Madrid, Spain, 1991.
93. Millán Vásquez de la Torre, M.G.; Hidalgo, L.A.; Arjona Fuentes, J.M. El oleoturismo: Una alternativa para preservar los paisajes del olivar y promover el desarrollo rural y regional de Andalucía (España). *Geografía Norte Grande* **2015**, *60*, 195–214. [CrossRef]

94. Mitcham, C. The concept of sustainable development: Its origins and ambivalence. *Technol. Soc.* **1995**, *17*, 311–326. [CrossRef]
95. Mokate, K.M. *Eficacia, Eficiencia, Equidad y Sostenibilidad: ¿Qué Queremos Decir?* Banco Interamericano de Desarrollo: Washington, DC, USA, 2001.
96. Mokate, K.M. *Evaluación Financiera de Proyectos de Inversión*, 2nd ed.; Alfaomega: Madrid, Spain, 2004.
97. Montensen, L.L. Teacher Education for Sustainability. I Global Change Education: The Scientific Foundation for Sustainability. *J. Sci. Educ. Technol.* **2000**, *9*, 27–36. [CrossRef]
98. Naredo, J.M. Sobre el origen, el uso y el contenido del término sostenible. *Cuadernos de Investigación Urbanística* **1996**, *41*, 7–18.
99. Nogales Naharro, M.A. Desarrollo rural y desarrollo sostenible. La sostenibilidad Ethics. *CIRIEC-España Revista de Economía Pública Social y Cooperativa* **2006**, *55*, 7–42.
100. O'Connor, M. Thermodynamique, complexité et codépendance écologique: La science de la joie et du deuil. *Revue Internationale de Systémique* **1994**, *8*, 397–423.
101. O'Connor, M. The "four spheres" framework for sustainability. *Ecol. Complex.* **2006**, *3*, 285–292. [CrossRef]
102. Organización de las Naciones Unidas. *Informe de Los Objetivos de Desarrollo Sostenible 2018*; Organización de las Naciones Unidas: New York, NY, USA, 2018.
103. O'Riordan, T. The Politics of Sustainability. In *Sustainable Environmental Management*; Turner, R.K., Ed.; Westview Press: Boulder, CO, USA, 1988; pp. 29–50.
104. Partridge, D.J. Activist capitalism and supply-chain citizenship: Producing ethical regimes and ready-to-wear clothes. *Curr. Anthropol.* **2011**, *52*, 97–111. [CrossRef]
105. Pearce, D. Foundations of an Ecological Economics. *Ecol. Model.* **1987**, *38*, 9–18. [CrossRef]
106. Pearce, J.; Albritton, S.; Grant, G.; Steed, G.; Zelenika, I. A new model for enabling innovation in appropriate technology for sustainable development. *Sustain. Sci. Pract. Policy* **2012**, *8*, 42–53. [CrossRef]
107. Pérez Rincón, M.A. Conceptualización sobre el Desarrollo Sostenible: Operacionalización del concepto para Colombia. *Punto de Vista* **2012**, *III*, 139–158. [CrossRef]
108. Pezzey, J. *Sustainability: An Interdisciplinary Guide*; Environmental Values; White Horse Press: Cambridge, UK, 1992; Volume 1, pp. 321–362. [CrossRef]
109. Pezzey, J.; Toman, M. The Economics of Sustainability. *Rev. J. Artic.* **2002**, *2*, 13–17. [CrossRef]
110. Pinillos, A.A.; Fernández Fernandez, J.L. De la RSC a la sostenibilidad corporativa: Una evolución necesaria para la creación de valor. *Bus. Rev.* **2011**, *207*, 5–21.
111. Plasencia Soler, J.A.; Marrero Delgado, F.; Bajo Sanjuán, A.M.; Nicado garcía, M. Modelos para evaluar la sostenibilidad de las organizaciones. *Estud. Gerenc.* **2018**, *34*, 63–73. [CrossRef]
112. Quiroga, R. *Indicadores de Sostenibilidad Ambiental y de Desarrollo Sostenible: Estado del Arte y Perspectivas*; Naciones Unidas: Santiago, Chile, 2001.
113. Rees, W.E. The ecology of sustainable development. *Ecologist* **1990**, *20*, 18–23.
114. Robilliard, P. *Los Ejes Centrales Para el Desarrollo de una Minería Sostenible*; CEPAL: Santiago, Chile, 2006.
115. Rodríguez-Ariza, L.; Frías Aceituno, J.V.; García Rubio, R. El consejo de administración y las memorias de sostenibilidad. *Revista de Contabilidad* **2014**, *17*, 5–16. [CrossRef]
116. Rodríguez, I.; Govea, H. El discurso del desarrollo sustentable en América Latina. *Revista Venezolana de Economía y Ciencias sociales* **2006**, *12*, 37–63.
117. Rojas, C.A.; Diaz, M.A.; Jaque Castillo, E. Sostenibilidad urbana. Tome: Una propuesta para evaluar los planes reguladores chilenos. *Urbano* **2009**, *11*, 26–35.
118. Rueda, G. Desarrollo alternativo y contabilidad pública: Hacia una línea de investigación. *Nómadas* **2002**, *16*, 195–203.
119. Schreiner, C.; Henriksen, E.K.; Kirkeby Hansen, P.J. Climate Education: Empowerin today's youth to meet tomorrow´s challenges. *Stud. Sci. Educ.* **2005**, *41*, 3–50. [CrossRef]
120. UICN; PNUMA; WWF. Cuidar la Tierra. In *Estrategia Para el Futuro de la Vida*; Hearthscan Edition: Gland, Switzerland, 1991.
121. Uribe Macías, M.E.; Vargas Moreno, O.A.; Merchán-Paredes, L. La responsabilidad social empresarial y la sostenibilidad, criterios habilitantes en gerencia de proyectos. *Entramado* **2018**, *14*, 52–63. [CrossRef]
122. Vásconez Carrasco, M.; Torres León, L. Minería en el Ecuador: Sostenibilidad y Licitud. *Estudios del Desarrollo Social Cuba y América Latina* **2018**, *6*, 48–50.
123. Vaitheeswaran, V. The great race. *Economics* **2002**, *364*, 3–5.
124. Vilches, A.; Gil Pérez, D. Ciencia de la sostenibilidad: Un nuevo campo de conocimientos al que la química y la educación química están contribuyendo. *Educación Química* **2013**, *24*, 199–206. [CrossRef]
125. Vilches, A.; Gil Pérez, D. Ciencia de la sostenibilidad: ¿Una nueva disciplina o un nuevo enfoque para todas las disciplinas? *Revista Iberoamericana de Educación* **2015**, *69*, 39–60. [CrossRef]
126. Willis, A. The role of the global reporting initiative's sustainability reporting guidelines in the social sustainability reporting guidelines in the social screening investments. *J. Bus. Ethics* **2003**, *43*, 233–237. [CrossRef]
127. World Bank. *Sustainable Development in a Dynamic World: Transforming Institutions, Growth and the Quality of Life*; Oxford University Press: New York, NY, USA, 2003.
128. World Commission on Environment and Development. *Report of the World Commission on Environment and Development: Our Common Future*; United Nations: New York, NY, USA, 1987.

mathematics

Article

Finance for the Environment: A Scientometrics Analysis of Green Finance

Rui Cai and Jianluan Guo *

School of Business, Central University of Finance and Economics, Beijing 100081, China; 2017110107@email.cufe.edu.cn
* Correspondence: guojianluan@email.cufe.edu.cn

Abstract: To protect environmental sustainability, organizations are moving their focus towards greening the business process. Similarly to any other business function, financial management has also turned to environmentally friendly activities. Green finance is a new financial pattern that integrates environmental protection and economic profits. This paper analyses the publications on green finance, and their intellectual structure and networking. The bibliometric data on green finance research have been extracted from the Scopus database. This study finds the most productive countries, universities, authors, journals, and most prolific publications in green finance, through examining the published works. Also, the study visualizes the intellectual network by mapping bibliographic coupling (BC) and co-citation. The study's essential contribution is the analysis of green finance developments and trajectories that can help scholars and practitioners to appreciate the trend and future studies.

Keywords: green finance; environmental finance; systematic review; future research; bibliometric analysis

1. Introduction

After the agrarian and industrial revolution, the negative impact of humans on the environment grew, specifically due to the incredible speed and scope of these new technologies, and the depletion of natural materials and energy [1,2]. Therefore, a concept such as Anthropocene was evolved to depict the state of human-dominated ecosystems. However, societal advancement and increasing the quality of human life are not accompanied by the growing Earth's capacity. Consequently, several environmental issues have begun to rise, which focus on ensuring human and societal advancement without affecting the ecosystem [3–5]. Some of these environmental issues pose a threat on a smaller scale, whilst others may have a more significant impact on the Earth's landscape and territories. For instance, unpredictable weather patterns, extreme heatwaves, ozone layer depletion, ocean acidification, deforestation, and the exploitation of the Earth's natural resources have resulted in unimaginable damage to habitable zones. In addition, they have caused great losses in biodiversity.

The gradual declension and corrosion of the environment has drawn the world's attention towards the human–environment relationship, and, in turn, environmental protection. Indeed, although humans have had an adverse impact on the environment, the human–nature relationship has changed; humans now wish to live in harmony with nature and employ sustainable development [6,7]. To cope with environmental issues, practitioners are moving their focus towards sustainable and green practices, and one such practice is green finance.

Over the past few years, the research on green finance has seen immense growth. The past literature aims to describe specific issues related to green finance, such as its scope, context, construct, antecedents, and consequences. Although a few attempts were made to conduct a literature review of green finance [8,9], these reviews did not provide a global perspective on green finance literature. A summarization of the research trends and trajectories would be helpful in order to develop future environmental-related studies in management and organizational strategy, and economic and financial policy-related areas. Thus, the authors suggest that a bibliometric analysis would be beneficial in profiling the existing and overall green finance research, as this method of analysis allows the objectivity of an area of research vis-à-vis a literature review [10,11]. It also helps steer through a bulk amount of bibliographic data [12,13]. Further, the bibliometric analysis serves to identify networking amongst academic groups, such as universities, countries, and journals, in a particular area of research [14–16].

This paper is grounded on the following two main objectives: firstly, it analyses the intellectual structure of green finance in terms of several indicators, such as the number of publications, number of citations, and networking (co-citation and bibliographic coupling); secondly, the paper suggests some future research agendas on green finance. The above targets can be further subdivided into the following phases: firstly, the paper describes a pattern between 2005 and 2020 in publication and citation; secondly, a global viewpoint evaluating the countries with the most significant number of documents and sources; thirdly, the top active universities and institutes are included in the list of the paper; fourthly, it recognizes the leading journals; fifthly, in terms of green finance studies, the most prolific authors are listed based on the number of publications and citations; sixthly, to see how these papers, countries, and authors are related, the study conducted bibliometric coupling and co-citation analysis; finally, we also propose a potential research agenda for green finance research, based on the data analysis.

This research makes two contributions. Firstly, it presents a basic overview of the key geographical regions, institutions, journals, authors, top-cited publications, and future research goals for green finance research. This study could also aid the researchers in determining the location. Further, it could help the universities whilst conducting their research. Besides, the study may also enable the policy makers to recognize the leading countries in green finance, to understand the optimal ecosystem, in order to guide their research and development efforts. Secondly, and most significantly, we propose a study plan that acknowledges theory-based opportunities and practical challenges that demand a future investigation.

2. Literature Review

Green finance is a relatively new field of finance. Economists and international organizations have failed to establish a precise definition or agree upon one unanimously. However, workable definitions have been developed by a variety of scholars, organizations, and governments [17]. An interesting variation in this is that specific organizations, rather than defining green finance, have coined the following phrase: a sustainable financial system [18]. Their tools and mechanisms, however, stay the same. According to the UNEP, a sustainable financial system incorporates the development of values and aids in dealing with financial assets, so that actual wealth may be used to meet the demands of an ecologically sustainable and inclusive economy over time. According to the Green Finance Study Group of the G20 [19], green financing is financing to promote the adoption of technologies that lower pollution. According to the People's Bank of China, green finance is a policy that refers to a set of policy and institutional arrangements aimed at attracting private money into green sectors, such as environmental preservation and energy conservation via financial services. In summary, green finance is defined as the "financing of investments that provide environmental benefits".

The notion of green financing is also different from the traditional banking approaches. It reflects the advantages of protecting the environment, by taking environmental risk management strategies and the sustainability of plans into account [20]. Furthermore, green financing strives to promote a green economy, in which the industries funded are anticipated to cut carbon emissions significantly.

Regarding adaptation to climate change concerns, the European Banking Federation takes a broader approach, stating that green finance is not restricted to simply environmental or climate change-related factors, therefore opening up opportunities for green insurance plans and green bonds [21]. As different as these definitions are, they all share some features in common, such as using capital for a broader and more sustainable purpose, benefitting the environment or reducing harm to it, managing risk to some extent, and framing policies and infrastructure to sustain the environment. To summarize, green finance is a subset of the self-sustaining financial system approach to addressing the challenges posed by climate change and the transition to a low-carbon society. Green finance may be defined as any financial investment that finances activities such as policy-making, insurance/risk solutions, bonds, or other commercial operations that have a much lower negative impact on the environment than the status quo, or positively influence the environment.

In addition, the green economy has three-fold benefits. Firstly, the development of green finance strengthens the role of corporate governance factors, and companies can realize the Pareto improvement of the natural ecological environment through the transformation to the green industry. Secondly, the green economy promotes environmental awareness and ensures that producers and consumers protect the environment by adapting to green energy and using biodegradable low-carbon products. Thirdly, green financial development can effectively optimize the supply structure of production factors, reduce overcapacity in traditional industries, and promote economic transformation and upgrading [22].

3. Methods

When scanning for academic publications on green finance, the Scopus database was used. The Scopus database is one of the largest peer-reviewed social science research repositories in the world. The repository is accessed and acknowledged for empirical and quantitative research [23–25]. We searched "Green Finance" in the article's title, abstract, keywords, and text, and we did not set any time limit for this search. The oldest paper dates to 2005, and a total of 172 papers were found.

When evaluating the data obtained, the bibliometric approach was used. Bibliometric is a library and information science research area, which uses quantitative techniques to analyze the bibliometric content [26,27]. This strategy helps identify and assess the overall trend of a particular topic, such as a journal, study area, or country [23,28,29]. Bibliometric studies have been used in literature to measure the relevance of a subject [30,31], the contribution of journals [32], educational institutes [29], and countries [28].

In this study, VOSviewer [33] and RStudio software were used to complete the analysis. They generate results in presentable forms of tables and graphics more effectively than other software applications [33]. Furthermore, these two sorts of software have been found to be immensely helpful when it comes to preparing and presenting the bibliometric coupling and co-citations. Bibliographic coupling is defined as "a relationship that occurs when two works reference the third work in their bibliographies," and two documents will be bibliographically coupled if both cite one or more documents in common. Co-citation is defined as "the frequency with which two documents are cited together by a third document" [34,35]. These are indications that the following probability exists: that the two works both treat a related subject matter.

4. Results

4.1. The Current State of Publishing

The search results of the Scopus database showed that there are 172 publications on green finance, and the oldest publication dates to 2005. A structure-based analysis of the publications (Table 1) shows that the number of publications in the last four years has increased significantly. The year 2020 has been the most productive year, with 62 journals.

This significant increase in the number of publications shows that acceptance of green finance as a separate area of research is taking off.

Table 1. The current state of publishing (line 154).

Year.	Number of Papers
2020	62
2019	44
2018	28
2017	13
2016	5
2015	6
2014	2
2013	3
2012	4
2011	2
2005	1

4.2. Most Productive Regions and Countries in Green Finance Research

Several countries are significantly contributing to green finance research. This section gives attention to the most productive countries between 2005 and 2020. Table 2 presents the results of the top 10 countries publishing green finance research. The ranking is based on the number of publications.

Table 2. The most productive countries (line 160).

No	Countries	NP	Citations	Citations/Paper
1	China	74	154	2.081
2	United Kingdom	16	165	10.313
3	United States	15	51	3.400
4	Germany	14	48	3.429
5	Japan	11	90	8.182
6	Canada	9	48	5.333
7	India	9	19	2.111
8	Italy	9	89	9.889
9	Russian Federation	8	16	2.000
10	Australia	7	59	8.429

The table shows that China is the most productive country, with 74 publications, indicating that Chinese academics are paying the most attention to green finance amongst all of the researchers. The UK is ranked 2nd with 16 journals, followed by the United States with 15 journals. However, China is not the only Asian country in the ranking table. The rest, Japan and India, also reached a significant position in green finance research, with 11 and 9 journals, respectively.

A bibliometric coupling (BC) is built to better understand networking amongst green finance-publishing countries. Bibliographic pairing happens when a third report is widely quoted in two documents. Concerning countries, it occurs when a paper from two separate countries references the third document in their publications, which highlights how nations use related literature in their journals.

The findings of the bibliometric coupling are depicted in Figure 1. Each circle represents a nation, and its size indicates the contribution, as follows: the larger the dot, the more outstanding the contribution. China is the world's most productive nation (consistent with Table 2). Further, China has the strongest bibliometric ties to other countries, followed by the United Kingdom and the United States.

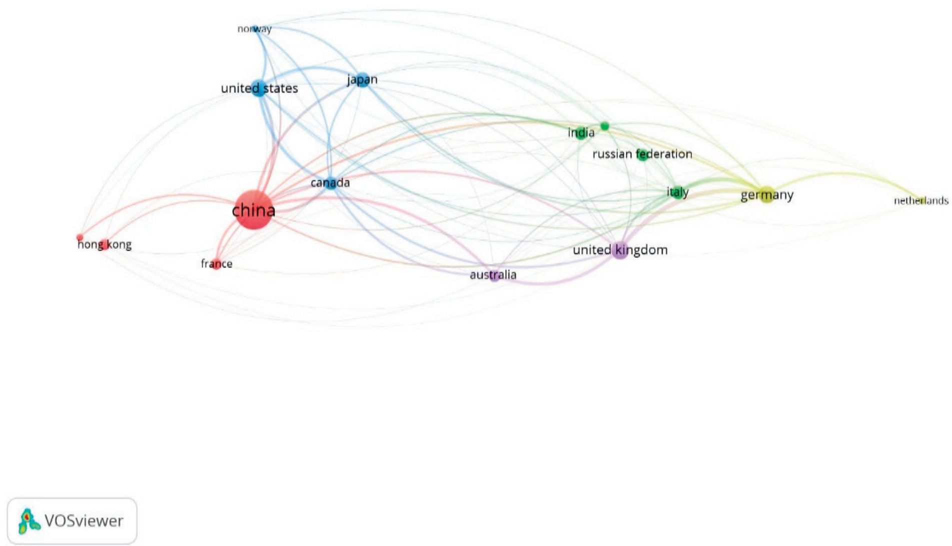

Figure 1. BC of countries (line 174).

Another fascinating question is how authors from various nations have interconnected with authors from other countries. A co-authorship analysis with governments accomplishes this. The co-authorship pattern is depicted in Figure 2, in terms of nations performing green finance research. Recall that we can determine a country's total number of publications and its critical ties with other nations through co-authorship.

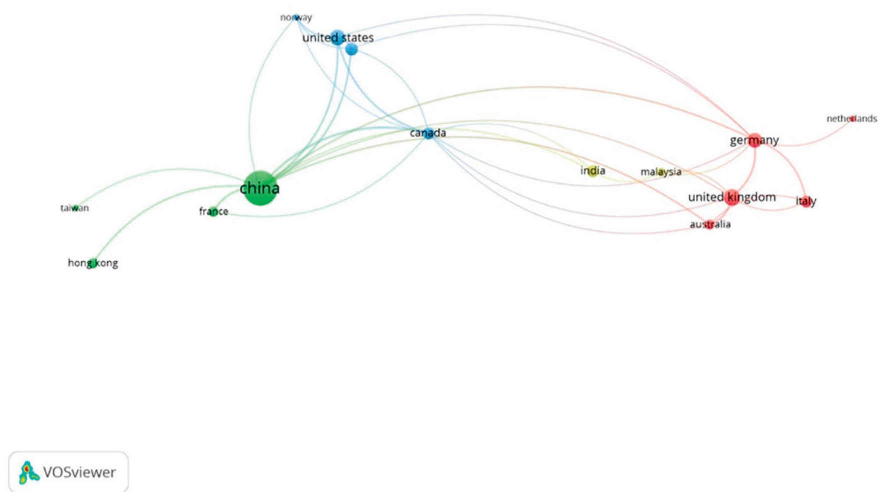

Figure 2. Co-authorship among countries (line 189).

Figure 2 represents the networking of co-authorship with different colors. For example, China is the most productive country, with the strongest co-authorship with other countries, meaning that the authors from other countries collaborate with Chinese authors.

4.3. Universities and Institutes with the Most Significant Number of Publications

Another significant aspect of the bibliometric analysis is to see which institute or university is the most productive. The results in Table 3 show that Hong Kong Polytechnic University is the most productive university, publishing five papers. Jinan University and Keio University are ranked 2nd and 3rd, with five and four papers, respectively. In the case of a tie, the authors are rated based on the number of citations their documents received.

Table 3. The most productive universities (line 196).

Universities/Institutes	TP	TC
Hong Kong Polytechnic University	5	21
Jinan University	5	6
Keio University	4	53
China University of Mining and Technology	4	18
Zhejiang University	3	0
Jiangsu Normal University	3	12
Ruhr-Universitat Bochum	3	19
University of Electronic Science and Technology of China	3	0
Kyushu University	3	20
Capital University of Economics and Business	3	0
Russian Academy of Sciences	3	0
University College London	3	4
Università degli Studi di Foggia	3	53
Deutsches Institut fur Entwicklungspolitik	3	12
Wuhan University	3	25
SOAS University of London	3	12
Unitelma Sapienza University	3	53
Primakov National Research Institute of World Economy and International Relations, Russian Academy of Sciences	3	4

4.4. Top Journals in Terms of Productivity

Another essential part of the bibliometric analysis is to analyze the most productive journals, which publish more research on green finance than others. Table 4 illustrates the journals publishing the research on green finance. The results are based on a minimum of 10 publications. Sustainability is the top journal, with 11 journals, followed by the Journal of Sustainable Finance and Investment, with 8 publications between 2005 and 2020. The Journal of Cleaner Production, with one paper less than the 2nd place publication, ranks 3rd.

Table 4. Top journals that publish green finance research (line 206).

No	Source Title	Papers	Citations	C/P
1	Sustainability	11	53	4.8
2	Journal of Sustainable Finance and Investment	8	8	1
3	Journal of Cleaner Production	7	35	5
4	Iop Conference Series: Earth and Environmental Science	6	2	0.3
5	Ecological Economics	4	57	14.3
6	Ekoloji	4	2	0.5
7	Modern China: Financial Cooperation for Solving Sustainability Challenges	4	0	0
8	Sustainable Energy and Green Finance for A Low-Carbon Economy: Perspectives From The Greater Bay Area Of China	4	2	0.5
9	World Economy and International Relations	4	15	3.8
10	Climate Policy	3	5	1.7
11	E3s Web of Conferences	3	0	0

The co-citation analysis of these top journals is an intriguing component of the bibliometric study. Co-citation happens when two papers from two distinct journals are cited jointly by a third document. For example, Figure 3 shows that the Energy Policy has the strongest co-citations amongst all of the journals, followed by the Journal of Cleaner Production and Sustainability. These facts show that the journals presented in Table 4 share similar literature, and that is why they have strong co-citations.

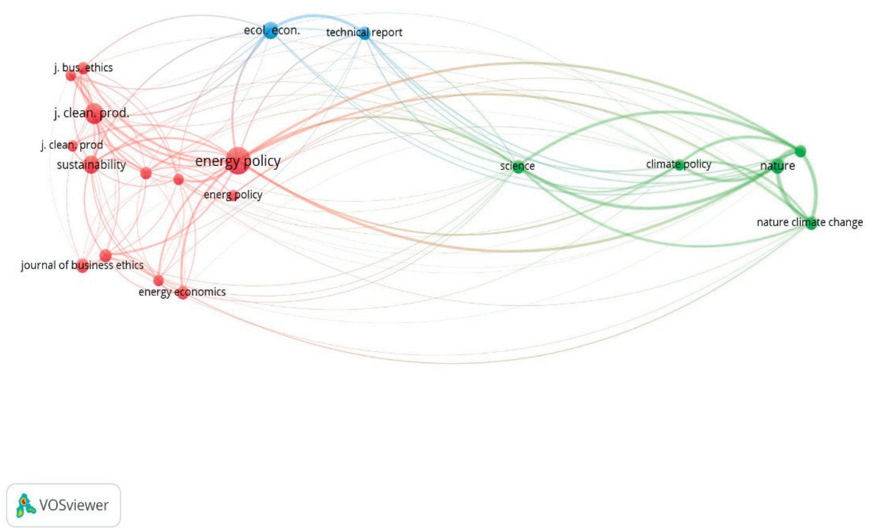

Figure 3. Co-citation of the top journals (line 215).

4.5. The Most Prolific Authors in Green Finance

See Table 5 to view the results of those who publish most frequently on green finance. It shows Falcone P.M., Ng A.W., Taghizadeh-Hesary F., and Yoshino N., each with four publications.

Table 5. The most productive authors in pro-environmental behavior (line 220).

No	Author	Papers	Citations	C/P
1	Falcone P.M.	4	53	13.3
2	Ng A.W.	4	21	5.3
3	Taghizadeh-Hesary F.	4	53	13.3
4	Yoshino N.	4	53	13.3
5	D'orazio P.	3	19	6.3
6	Fu J.	3	2	0.7
7	Liu C.	3	3	1
8	Managi S.	3	28	9.3
9	Sica E.	3	53	17.7
10	Volz U.	3	12	4
11	Wang Y.	3	30	10
12	Xia Y.	3	12	4

4.6. The Most-Cited Publication

Another critical factor when it comes to evaluating the output of authors, journals, and papers is the most-cited source or reference. Therefore, we listed the most-cited publication in this section. We relied on the data from the Scopus database. As stated in the Method section, we searched "Green Finance" in the article's title, abstract, keywords, and text. This process provided us with a list of all those publications with green finance and the related keywords mentioned above in their title, abstract, or keywords. Table 6 presents the list of publications that received more than 10 citations.

4.7. Three-Field Plot

To determine the author's area of specialization in green finance, we conducted a three-field plot analysis. This study establishes a link between authors and the keywords they included in their green finance papers. In Figure 4, the authors' names are listed on the left, the keywords used in the green finance paper are listed in the middle, and the authors' countries are shown on the right. We may easily deduce the author's field of study from the direction of the links. For example, Liu C mainly used green finance, green economy, and sustainable development as keywords, so his research mainly concerns green economic problems in sustainable development from a financial perspective. Similarly, Ng A.W. mainly wrote on green finance, sustainability, and green bonds, which indicates his research was specific to green bond instruments.

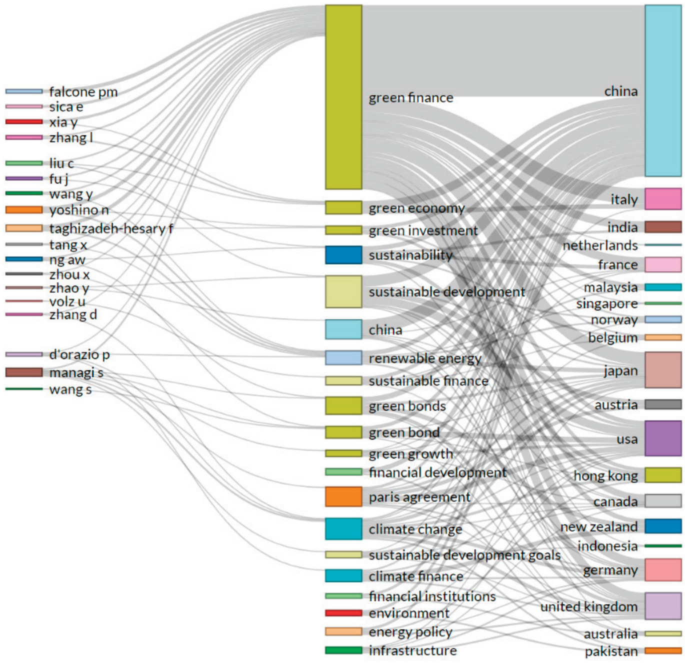

Figure 4. Three-field plot representing authors' area of research (line 242).

4.8. Co-Occurrence of Authors' Keywords

This section analyses the content further, in terms of keywords, which is a valuable tool for locating relevant literature and trends. In this phase of the investigation, VOSviewer software is used to discuss the co-occurrence of the authors' keywords to explain the features of publications in green finance, and the results are displayed in Figure 5. Note that the keywords analyzed here are based on the text content of keywords in papers. Also, the co-occurrence implies the frequency that a keyword appears with other keywords. In this research, keyword analyses are utilized to find trends in the subjects examined.

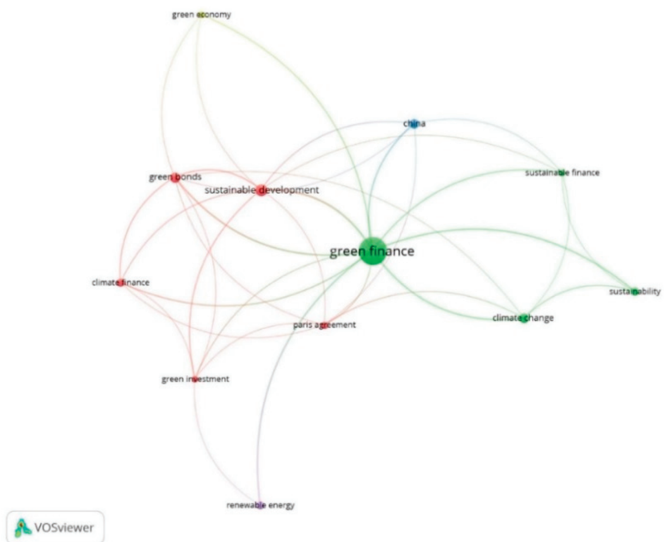

Figure 5. Co-occurrence of authors' keywords (line 250).

Table 6. The most-cited publication (line 229).

Authors	Title	Year	Source Title	Number of Citations
Dafermos Y., Nikolaidi M., Galanis G.	A Stock-Flow-Fund Ecological Macroeconomic Model	2017	Ecological Economics	38
Falcone P.M., Morone P., Sica E.	Greening of the Financial System and Fuelling A Sustainability Transition: A Discursive Approach to Assess Landscape Pressures on The Italian Financial System	2018	Technological Forecasting and Social Change	33
Taghizadeh-Hesary F., Yoshino N.	The Way to Induce Private Participation in Green Finance and Investment	2019	Finance Research Letters	31
Wang Y., Zhi Q.	The Role of Green Finance in Environmental Protection: Two Aspects of Market Mechanism and Policies	2016	Energy Procedia	30
Clark R., Reed J., Sunderland T.	Bridging Funding Gaps for Climate and Sustainable Development: Pitfalls, Progress and Potential of Private Finance	2018	Land Use Policy	23
Falcone P.M., Sica E.	Assessing the Opportunities and Challenges of Green Finance in Italy: An Analysis of The Biomass Production Sector	2019	Sustainability (Switzerland)	20
Ng A.W.	From Sustainability Accounting to A Green Financing System: Institutional Legitimacy and Market Heterogeneity in A Global Financial Centre	2018	Journal of Cleaner Production	20
Zhang D., Zhang Z., Managi S.	A Bibliometric Analysis on Green Finance: Current Status, Development, and Future Directions	2019	Finance Research Letters	19
Zahar A., Peel J., Godden L.	Australian Climate Law in Global Context	2012	Australian Climate Law in Global Context	18
Cui Y., Geobey S., Weber O., Lin H.	The Impact of Green Lending on Credit Risk in China	2018	Sustainability (Switzerland)	15
Bai Y., Faure M., Liu J.	The Role of China's Banking Sector in Providing Green Finance	2013	Duke Environmental Law and Policy Forum	15
Qin J., Zhao Y., Xia L.	Carbon Emission Reduction with Capital Constraint Under Greening Financing and Cost Sharing Contract	2018	International Journal of Environmental Research and Public Health	14
D'Orazio P., Popoyan L.	Fostering Green Investments and Tackling Climate-Related Financial Risks: Which Role for Macroprudential Policies?	2019	Ecological Economics	13
Yoshino N., Taghizadeh-Hesary F., Nakahigashi M.	Modelling The Social Funding and Spill-Over Tax for Addressing The Green Energy Financing Gap	2019	Economic Modelling	13
Meyer C.A.	Public-Nonprofit Partnerships and North-South Green Finance	1997	Journal of Environment and Development	13

Table 6. *Cont.*

Authors	Title	Year	Source Title	Number of Citations
Jin J., Han L.	Assessment of Chinese Green Funds: Performance and Industry Allocation	2018	Journal of Cleaner Production	12
Soundarrajan P., Vivek N.	Green Finance for Sustainable Green Economic Growth in India	2016	Agricultural Economics (Czech Republic)	12
Braungardt S., van den Bergh J., Dunlop T.	Fossil Fuel Divestment and Climate Change: Reviewing Contested Arguments	2019	Energy Research and Social Science	11
He L., Liu R., Zhong Z., Wang D., Xia Y.	Can Green Financial Development Promote Renewable Energy Investment Efficiency? A Consideration of Bank Credit	2019	Renewable Energy	10
Glomsrød S., Wei T.	Business as Unusual: The Implications of Fossil Divestment and Green Bonds for Financial Flows, Economic Growth and Energy Market	2018	Energy for Sustainable Development	10
Xiong L., Qi S.	Financial Development and Carbon Emissions in Chinese Provinces: A Spatial Panel Data Analysis	2018	Singapore Economic Review	10
Porfir'ev B.N.	Green Trends in The Global Financial System	2016	World Economy and International Relations	10
Mathews J.A., Kidney S.	Financing Climate-Friendly Energy Development Through Bonds	2012	Development Southern Africa	10

Current scholars can use keyword analysis to find unidentified themes and domains in the literature. Of course, it is impossible to anticipate which subjects will receive more attention in the future. Nevertheless, based on the co-occurrence of the terms, we believe that the current trend will continue.

4.9. The Thematic Map: A Co-Word Analysis

To draw a thematic map (Figure 6) of the keywords of green finance, we used the Biblioshiny package of RStudio, with a minimum frequency of five per thousand words. This technique helps to map the main themes of the journal. Based on the suggestion proposed by Cobo et al. [36], this strategic map plotted the main themes into four quadrants, according to their density and centrality ranks along the *X* and *Y* axes. Centrality is a measure of the interaction of a network with other networks, and it is the measure of the importance of theme in the whole research; in this case, green finance [36]. In comparison, density identifies the degree of development of themes based on the internal strengths of the network.

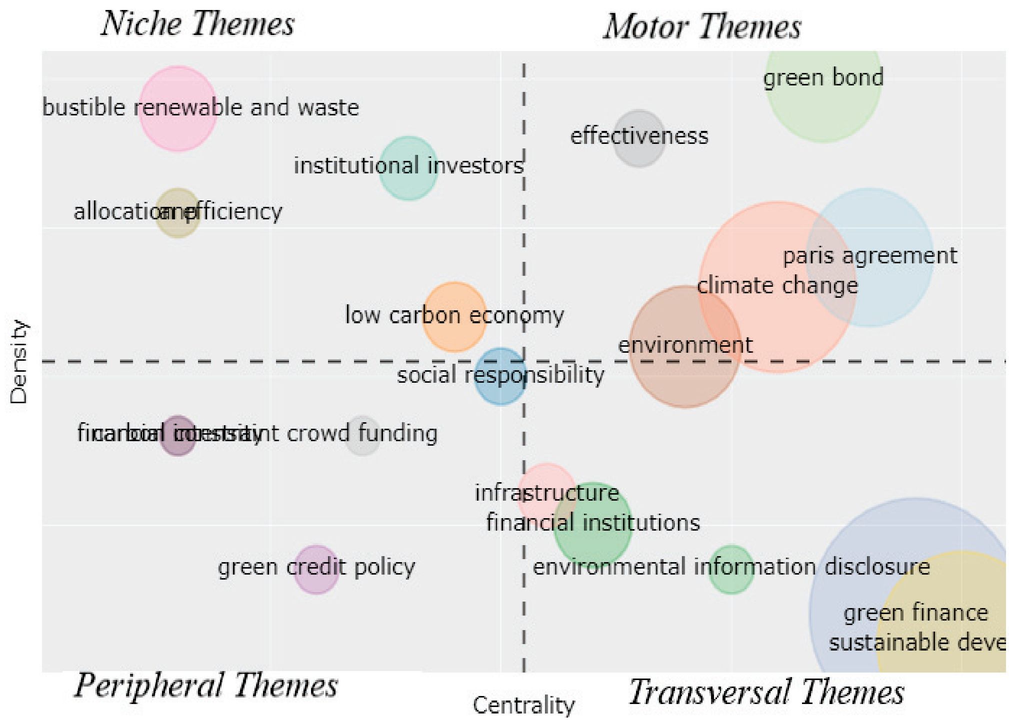

Figure 6. Thematic map of the journal (line 294).

4.9.1. Motor Themes

These themes are well developed and constitute the critical structure of the research in the journal. Moreover, these themes share high centrality and density. This quadrant contains climate change, the Paris Agreement, effectiveness, green bond, and environment. These facts show that the journal's major developed research areas are based on environmental and global climate issues, and green bond is the most advanced theme.

4.9.2. Niche Themes

Niche themes are very specialized and well-developed themes of a research area or journal. According to the results presented in Figure 6, the journal's well-developed and specialized research areas are institutional investors, allocations efficiency, and low-carbon economy. The most used concept is bustible renewable and waste.

4.9.3. Peripheral Themes

The third quadrant consists of the themes that are either emerging or declining in the journal. These themes are characterized by low centrality and low density. This quadrant includes social responsibility, crowdfunding, green credit policy, financial industry, and carbon density (they overlap in the image). From the Figure 6, we can see that social responsibility is moving towards a higher centrality, which shows its more universal significance when compared with the rest of the quadrants in the journal.

4.9.4. Transversal Themes

The fourth quadrant is characterized as low density and high centrality. These themes are essential for the research, but are not well developed. Interestingly, specific themes, such as green finance and sustainable development, are further away from the motor themes and peripheral themes. These themes have a lower density, but higher centrality, than the other themes present in this quadrant.

5. Topic Trends and Some Future Research Suggestions

In addition to the keyword co-occurrence analysis, it is also of great interest to understand the popularity of different topics and themes during the past 15 years, from the time window. On the one hand, from Figure 7, we can see that during the early years of green finance publication (2005–2014), finance was the only area of research in the journal. On the other hand, from 2015 to 2018, research on corporate social responsibility, green growth, low-carbon economy, and sustainability, which studies environment protections from an economic perspective, has been in fashion. Until 2019, green finance, green bonds, and green investment became independent topics amongst researchers.

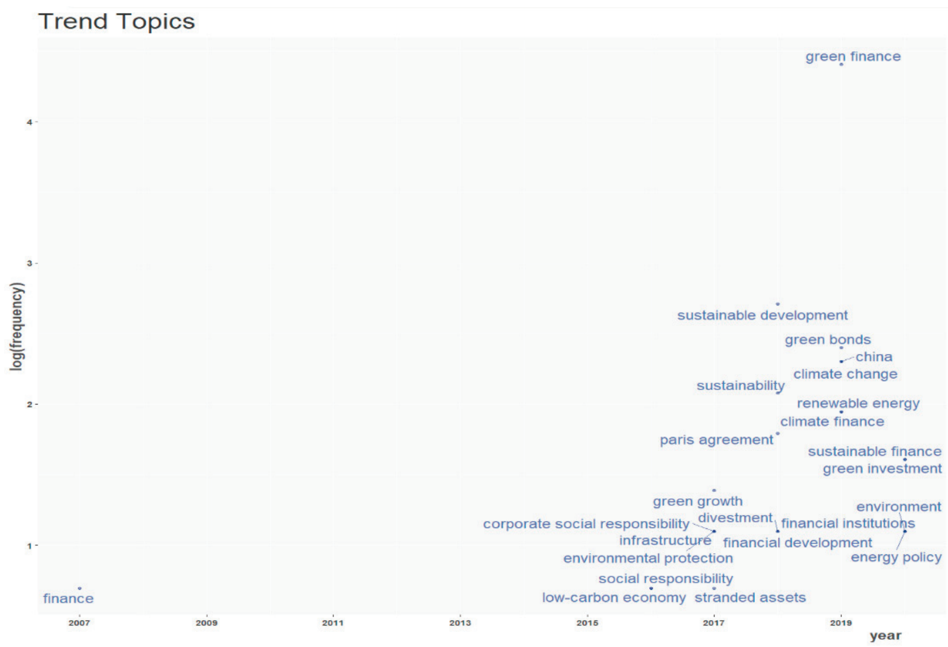

Figure 7. Trends of the topic in the journal (line 317).

The thesis established in this paper had the bibliometric analysis of green finance as its core focus. That is a new debate, and there are still several possibilities for fresh research prospects, including the holes found in this report. Current environmental problems lead to tremendous pressure on businesses and the public sector to participate in regular debates, mainly on financial viability, which profoundly affects corporate behavior.

Since firms concentrate on growing their financial profits, sustainable activities in the corporate culture are necessary for them to be implemented. Therefore, a holistic vision of green finance is critical to tackling this issue. The thesis discusses this vital question in this context, adopting an objective bibliometric method and obtaining results from various dimensions.

This essay was designed to present the key topics in the area. Therefore, the holes can be considered as an incentive for future research, looking at the content presented in this discussion platform until 2020. The trends were obtained by the study. After reviewing it in association with previous results, we were able to see that a research agenda already remains. This needs to be further explored in new studies in economics, entrepreneurship, and sustainability. Its great value and fast-growing relevance have been demonstrated by green finance. With growing curiosity in the recent literature, but a lack of coverage from mainstream economics and finance journals, a void has been generated that offers researchers prospects for potential advances in at least three directions.

Firstly, given that green finance is fundamentally a finance topic, there is an immediate need to research the problems of green finance from the viewpoint of finance and the use of financial techniques. In conventional finance papers, green bonds, green risk management, and green governance may be of concern.

Secondly, further research from developed countries' viewpoints on green finance issues will be beneficial for regulators and decision-makers to coordinate diverse policy priorities and develop well-defined policy targets. More resources will be offered to scholars from developing countries with specific knowledge advantages, and more international cooperation between developing and developed countries is expected.

Thirdly, it should be noted that a remarkable contrast between the problems of green finance and traditional finance is that the former is primarily guided by regulation.

Thus, with the increasingly evolving international economic and political climate, new problems in this area are expected to arise. In addition, to the best of our understanding, there is still no systematic analysis of relevant literature available. It is certainly worth investigating the findings of our bibliometric research and making comparisons.

6. Limitations

Despite this study's contribution, it experiences some shortcomings. Firstly, the study is based on bibliometric data derived from the Scopus database, but the shortcomings of this database also weaken the universality of the study. In addition, there might be several papers and journals that are indexed in the Web of Sciences, but are not indexed in the Scopus. Those are exempt from this study. We also recommend that future research extracts information from other databases, such as the Science Site and the host, EBSCO. Secondly, the research performed a retrospective review from 2005 to 2020, in terms of visual analysis. Therefore, we recommend that future studies complete some temporal analyses to see how patterns in green finance literature grow over different years.

Author Contributions: Conceptualization, R.C.; methodology, R.C.; validation, J.G.; formal analysis, R.C.; investigation, R.C.; resources, R.C.; data curation, R.C.; writing—original draft preparation, R.C.; writing—review and editing, J.G.; visualization, R.C.; supervision, J.G.; project administration, J.G.; funding acquisition, J.G. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by humanities and social science research, and the Planning Foundation of the Ministry of Education of China, grant number 19YJA630022, and The APC was funded by Beijing Natural Science Foundation Project, grant number 9192002.

Data Availability Statement: All data involved in this study were obtained from the Scopus database.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Wang, S.; Wang, J.; Li, J.; Yang, F. Do motivations contribute to local residents' engagement in pro-environmental behaviors? Resident-destination relationship and pro-environmental climate perspective. *J. Sustain. Tour.* **2020**, *28*, 834–852. [CrossRef]

2. Ansari, N.; Farrukh, M.; Raza, A. Green human resource management and employees pro-environmental behaviours: Examining the underlying mechanism. *Corp. Soc. Responsib. Environ. Manag.* **2020**, *28*, 229–238. [CrossRef]
3. Whitburn, J.; Linklater, W.; Abrahamse, W. Meta-analysis of human connection to nature and proenvironmental behavior. *Conserv. Biol.* **2020**, *34*, 180–193. [CrossRef]
4. Ture, R.S.; Ganesh, M.P. Pro-environmental behaviours at workplace: An empirical study in Indian manufacturing organizations. *Benchmarking Int. J.* **2018**, *25*, 3743–3766. [CrossRef]
5. Raza, A.; Farrukh, M.; Iqbal, M.K.; Farhan, M.; Wu, Y. Corporate social responsibility and employees' voluntary pro-environmental behavior: The role of organizational pride and employee engagement. *Corp. Soc. Responsib. Environ. Manag.* **2021**. [CrossRef]
6. Yong, J.Y.; Yusliza, M.Y.; Fawehinmi, O.O. Green human resource management: A systematic literature review from 2007 to 2019. *Benchmarking Int. J.* **2019**, *27*, 2005–2027. [CrossRef]
7. Clement, C.; Lochan, D.R. Green competencies: Insights and recommendations from a systematic literature review. *Benchmarking Int. J.* **2020**. [CrossRef]
8. Li, D.; Zhao, L.; Ma, S.; Shao, S.; Zhang, L. What influences an individual's pro-environmental behavior? A literature review. *Resour. Conserv. Recycl.* **2019**, *146*, 28–34. [CrossRef]
9. Inoue, Y.; Alfaro-Barrantes, P. Pro-environmental behavior in the workplace: A review of empirical studies and directions for future research. *Bus. Soc. Rev.* **2015**, *120*, 137–160. [CrossRef]
10. Farrukh, M.; Raza, A.; Meng, F.; Wu, Y. CMS AT 13: A retrospective of the journey. *Chin. Manag. Stud.* **2020**, 1–29. [CrossRef]
11. Gu, Z.; Meng, F.; Farrukh, M. Mapping the research on knowledge transfer: A scientometrics approach. *IEEE Access* **2021**, *9*, 34647–34659. [CrossRef]
12. Gao, P.; Meng, F.; Mata, M.N.; Martins, J.M.; Iqbal, S.; Correia, A.B.; Dantas, R.M.; Waheed, A.; Rita, J.X.; Farrukh, M. Trends and future research in electronic marketing: A bibliometric analysis of twenty years. *J. Theor. Appl. Electron. Commer. Res.* **2021**, *16*, 94. [CrossRef]
13. Nawaz, K.; Saeed, H.A.; Sajeel, T.A. Covid-19 and the state of research from the perspective of psychology. *Int. J. Bus. Psychol.* **2020**, *2*, 35–44.
14. Farrukh, M.; Raza, A.; Javed, S.; Lee, J.W.C. Twenty years of green innovation research: Trends and way forward. *World J. Entrep. Manag. Sustain. Dev.* **2021**. [CrossRef]
15. Ahmed, S.I.; Farrukh, M.; Yihua, W.; Trunk, N. Human systems management: A retrospective of 40 years. *Hum. Syst. Manag.* **2021**, *41*, 15–30.
16. Nawaz, K.; Aslam, T.; Saeed, H.A. A bibliometric analysis of international journal of sports marketing & sponsorship. *Int. J. Bus. Psychol.* **2020**, *2*, 45–60.
17. Labatt, S.; White, R.R. *Environmental Finance: A Guide to Environmental Risk Assessment and Financial Products*; John Wiley & Sons: New York, NY, USA, 2003.
18. Hira, T.K. Promoting sustainable financial behaviour: Implications for education and research. *Int. J. Consum. Stud.* **2012**, *36*, 502–507. [CrossRef]
19. Berensmann, K.; Volz, U.; Alloisio, I.; Bak, C.; Bhattacharya, A.; Leipold, G.; Schindler, H.; MacDonald, L.; Huifang, T.; Yang, Q. Fostering Sustainable Global Growth through Green Finance—What Role for the G20. *T20 Task Force Clim. Policy Financ.* **2017**. Available online: https://collaboration.worldbank.org/content/usergenerated/asi/cloud/attachments/sites/collaboration-for-development/en/groups/green-finance-community-of-practice/documents/jcr:content/content/primary/blog/green_finance_educat-LVNC/Climate_Green-Finance_V2.pdf (accessed on 8 June 2021).
20. Jeucken, M. *Sustainable Finance and Banking: The Financial Sector and The Future of The Planet*; Earthscan: London, UK, 2010.
21. European Banking Federation. *Towards A Green Finance Framework*; Brussels, Belgium; Frankfurt, Germany, 2017; Available online: https://www.ebf.eu/wp-content/uploads/2017/09/Geen-finance-complete.pdf (accessed on 8 June 2021).
22. Bergset, L. The rationality and irrationality of financing green start-ups. *Adm. Sci.* **2015**, *5*, 260–285. [CrossRef]
23. Farrukh, M.; Meng, F.; Wu, Y.; Nawaz, K. Twenty-eight years of business strategy and the environment research: A bibliometric analysis. *Bus. Strateg. Environ.* **2020**, *29*, 2572–2582. [CrossRef]
24. Farrukh, M.; Shahzad, I.A.; Meng, F.; Wu, Y.; Raza, A. Three decades of research in the technology analysis & strategic management: A bibliometrics analysis. *Technol. Anal. Strateg. Manag.* **2020**, 1–17. [CrossRef]
25. Wu, Y.; Farrukh, M.; Raza, A.; Meng, F.; Alam, I. Framing the evolution of the corporate social responsibility and environmental management journal. *Corp. Soc. Responsib. Environ. Manag.* **2021**. [CrossRef]
26. Rose, R.; Hölzle, K.; Björk, J. More than a quarter century of creativity and innovation management: The journal's characteristics, evolution, and a look ahead. *Create. Innov. Manag.* **2020**, *29*, 5–20. [CrossRef]
27. Farrukh, M.; Meng, F.; Raza, A. Twenty-seven years of sustainable development journal: A bibliometric analysis. *Sustain. Dev.* **2020**, *28*, 1725–1737. [CrossRef]
28. Bonilla, C.A.; Merigó, J.M.; Torres-Abad, C. Economics in Latin America: A bibliometric analysis. *Scientometrics* **2015**, *105*, 1239–1252. [CrossRef]
29. Martínez-López, F.J.; Merigó, J.M.; Valenzuela-Fernández, L.; Nicolás, C. Fifty years of the European journal of marketing: A bibliometric analysis. *Eur. J. Mark.* **2018**, *52*, 439–468. [CrossRef]
30. Laengle, S.; Merigó, J.M.; Miranda, J.; Słowiński, R.; Bomze, I.; Borgonovo, E.; Dyson, R.G.; Oliveira, J.F.; Teunter, R. Forty years of the European journal of operational research: A bibliometric overview. *Eur. J. Oper. Res.* **2017**, *262*, 803–816. [CrossRef]

31. Garrigos-Simon, F.J.; Botella-Carrubi, M.D.; Gonzalez-Cruz, T.F. Social capital, human capital, and sustainability: A bibliometric and visualization analysis. *Sustainability* **2018**, *10*, 4751. [CrossRef]
32. Amiguet, L.; Lafuente, A.M.; Kydland, F.E.; Lindahl, J.M.M. One hundred and twenty-five years of the journal of political economy: A bibliometric overview. *J. Polit. Econ.* **2017**, *125*, 105–110.
33. Van Eck, N.; Waltman, L. Software survey: VOSviewer, a computer program for bibliometric mapping. *Scientometrics* **2010**, *84*, 523–538. [CrossRef]
34. Small, H. Co-citation in the scientific literature: A new measure of the relationship between two documents. *J. Am. Soc. Inf. Sci.* **1973**, *24*, 265–269. [CrossRef]
35. Baraibar-diez, E.; Luna, M.; Odriozola, D.; Llorente, I. Mapping social impact: A bibliometric analysis. *Sustainability* **2020**, *12*, 9389. [CrossRef]
36. Cobo, M.J.; López-Herrera, A.G.; Herrera-Viedma, E.; Herrera, F. An approach for detecting, quantifying, and visualizing the evolution of a research field: A practical application to the fuzzy sets theory field. *J. Informetr.* **2011**, *5*, 146–166. [CrossRef]

mathematics

Article

Instruments and Methods for Identifying Indicators of a Digital Entrepreneurial System

Jelena Raut [1], Đorđe Ćelić [1,*], Branislav Dudić [2,3,*], Jelena Ćulibrk [1] and Darko Stefanović [1]

[1] Faculty of Technical Sciences, University of Novi Sad, 21000 Novi Sad, Serbia; jelenaraut@uns.ac.rs (J.R.); jculibrk@uns.ac.rs (J.Ć.); darko.stefanovic@uns.ac.rs (D.S.)
[2] Faculty of Management, Comenius University in Bratislava, 82005 Bratislava, Slovakia
[3] Faculty of Economics and Engineering Management, University Business Academy, 21000 Novi Sad, Serbia
* Correspondence: celic@uns.ac.rs (Đ.Ć.); branislav.dudic@fm.uniba.sk (B.D.)

Abstract: Entrepreneurial ecosystems are the main driver of the widespread trend of digitalization, and they open opportunities for the advancement of the digital economy. The digital economy makes its progress through innovative enterprises that can ensure global progress. In order to effectively use the opportunities that open up the process of digitalization, information is needed on how much the Republic of Serbia is able to support the process of discovering entrepreneurship, which is stimulated by digitalization, which is the subject of this paper. The aim of this paper is to analyze the digital entrepreneurial system of the Republic of Serbia, as well as to identify indicators that hinder the development of this system, using appropriate instruments and methods that will be presented in detail in the paper. The results have demonstrated that the starting point for improvement of the digital entrepreneurial system is in the field of finance, with a particular focus on companies in the startup and stand-up phases. Furthermore, a comparative analysis will showcase the digital entrepreneurial system of the Republic of Serbia and the member states of the European Union, where it will be seen that the digital entrepreneurial system of the Republic of Serbia is lagging behind the member states of the European Union in its growth and development. The results will serve as the starting point for policymakers to improve the process of digitalization and the digital entrepreneurial system as a whole. The results show the starting point for the improvement of entrepreneurship in the Republic of Serbia, that is, how small, and medium-sized enterprises can be encouraged on the path to their successful management.

Keywords: digital entrepreneurial ecosystem; digital technologies; digital entrepreneurship; digital systems; innovations; small and medium-sized enterprises

1. Introduction

Digitalization is a process to which a lot of scientific research is dedicated, and it is essential because it shapes the nature and location of entrepreneurial opportunities. The development of digitalization has encouraged the global adoption of new organizational innovations, which will support entrepreneurial opportunities The table must appear after the citation in the main text1]. With the development of digitalization, a new type of regional agglomeration of economic activities is developing, identified as an entrepreneurial ecosystem [1].

Digitalization is a process that is of defining importance for entrepreneurial ecosystems. Digitalization enables digital technologies and digital infrastructures to be woven into the economy and society [2]. Two crucial characteristics of digital technologies and infrastructures stand out, which help to explain the transformational impact they have on innovation and entrepreneurship. The first significant characteristic is that digital technologies and infrastructures have a general purpose. They can be applied in almost all sectors, functions, and activities, potentially transforming them [3]. The second important characteristic is that digital technologies and digital infrastructures disclose entirely new ways of organizing. When these two characteristics are combined, they make digital technologies and digital infrastructures more powerful, at the same time facilitating innovation of business models [4]. Thus, it can be concluded that digitalization is a process that conditions certain changes in the economy by enabling the execution of entirely new processes or the execution of existing processes, in a significantly different way, which in most cases is more efficient than the former. It is precisely these dynamics that are digitally enabled and that emerge with the discovery of new business models that make entrepreneurial ventures become a central driver of the development of the digital economy and allow for the unlocking of the full productivity potential [4].

Digital entrepreneurship is a category of entrepreneurship, which is accomplished by taking advantage of digital technologies and digital infrastructures. Digital entrepreneurship can be defined as the digitalization of one or more predefined business models by entrepreneurs in terms of distribution and workplace [5]. Furthermore, digital entrepreneurship is a subcategory of entrepreneurship, which should be physically digitalized in a traditional organization [6]. Digital entrepreneurship is an attempt by the current entrepreneurial environment to adapt to the needs of the digital age in the context of developmental and changing conditions [7]. Digital entrepreneurship is also a search for opportunities through digital media and other information and communication technologies [8]. Digital entrepreneurship can also be defined as the encounter of digital information and technology with entrepreneurs. An entrepreneur is a person who uses the internet to create commercial opportunities, disseminate information, cooperate with customers or partners [9]. Due to its great importance, digitalization is included in research and encompasses the conditions of digitalization that are binding throughout the country.

The components that make up the entrepreneurial ecosystem greatly influence the shaping of digital entrepreneurship. Entrepreneurial ecosystems encourage entrepreneurs to generate ideas, identify and allocate resources, and take advantage of the opportunities provided by the digital marketplace, where they will gather information and create legitimacy for innovation [5].

The entrepreneurial ecosystem has emerged as an entirely new cluster type to support the digital dynamics of discovering a new business model [10]. The entrepreneurial ecosystem represents the community of all stakeholders who specialize in facilitating the improvement of business models. A brief literature review suggests a view of entrepreneurial ecosystems as a new cluster type that characterizes the digital economy.

The latest evolution in entrepreneurial policy emphasizes a holistic and multifaceted view of entrepreneurship, where the entrepreneurial ecosystem emerges as a result. A holistic approach to entrepreneurship focuses on the individual, personality, and behavior, contributing to a better understanding of entrepreneurial activity [11–13]. Furthermore, a holistic approach advocates the research of entrepreneurial activity, not focusing on entrepreneurial activities in isolation, but as individual behavior of entrepreneurs, which is embedded in the local context [14–16]. The focus on entrepreneurial ecosystems draws attention to policies and initiatives for entrepreneurship throughout the life cycle of an entrepreneurial venture. Support for entrepreneurship must be viewed in a broader regional context [17]. As stated in the REDI (The Regional Entrepreneurship and Development Index) Measuring Regional Entrepreneurship Final Report, the concepts of the entrepreneurial ecosystem and the entrepreneurship system are very similar. However, the difference is that entrepreneurial ecosystems focus on policies and initiatives to support entrepreneurship from policy perspectives. In contrast, entrepreneurship systems focus on the entrepreneurial dynamics that drive productivity growth in the region [12].

The entrepreneurial ecosystem can be defined as a new way of contextualizing more complex social systems [18]. A system is an organized set of interactive and interdependent systems functioning as a whole to achieve a specific purpose. At the same time, an ecosystem is a purposeful network of dynamic and interactive systems and subsystems that have a variable set of privacy in today’s context [19]. The entrepreneurial ecosystem consists of subsystems integrated into systems, which can be optimized for system performance at the ecosystem level [20]. According to the authors of the Global Entrepreneurship Index powered by GEDI, the concept of the entrepreneurship system is based on three premises, which provide a starting point for the analysis of the entrepreneurial ecosystem [20]. The first premise is based on the fact that entrepreneurship is based on actions taken and led by individuals, based on incentives; the second premise argues that individual action is affected by framework conditions; while the third premise claims that ecosystems are complex, multilayered structures in which many elements interact and affect system performance.

Entrepreneurial ecosystems can also be defined at the socio-economic level, where they have the properties of self-organization, scalability, and sustainability as "dynamic institutional policy, built-in interaction between entrepreneurial attitudes, abilities, and aspirations by individuals, which drive resource allocation through creating and launching new ventures" [21]. Entrepreneurial ecosystems are complex socio-economic structures, which revive the actions of individual levels, where a lot of knowledge is relevant to entrepreneurial activity and is built into the structures of ecosystems, and require action at the individual level. Entrepreneurial ecosystems are also defined as resource allocation systems, which facilitate the allocation of resources according to productive use [22]. The widespread trend of digitalization plays a significant role in the advancement of entrepreneurial ecosystems, which opens the possibility for every organization to reconsider value creation activities in the economy through business model innovations. This very feature enables entrepreneurial ecosystems to become a key driver of progress towards a digital economy.

Digitalization is a process that reorganizes the business activities of a company and the entire society around digital technologies and digital infrastructure. The entrepreneurial ecosystem represents a new form of regional joining of entrepreneurial activities, which supports the process of discovering new radical business models, which are a challenge for existing companies [4]. This is a unique entrepreneurial challenge that, in a way, conditions companies to adopt new and more efficient practices. That is why entrepreneurial ecosystems are significant for improving the productivity potential of the digital economy.

The authors of this paper focused on the analysis of the digital entrepreneurial system of the Republic of Serbia, based on the report The European Index of Digital Entrepreneurial Systems (EIDES), to make recommendations for more effective policies for progress towards the digital economy, based on research pertaining to the degree of Serbia's ability to support the process of discovering entrepreneurship encouraged by digitalization. Within the EIDES report, the Republic of Serbia was not included in the research.

In every digital entrepreneurial system, there are stakeholders, which differ from one digital entrepreneurial system to another, as well as other elements that create certain results within the system, i.e., innovative companies that compete with each other with their digitally improved business models. These companies provide an "economic service" through which they improve the digital economy by reorganizing their businesses and creating new values. This is another reason the authors decided to lay the foundation of the research based on the EIDES report, because it was designed to cover all aspects of this fundamental dynamic.

At the very end, the authors will conclude which elements of the digital entrepreneurial system of the Republic of Serbia are the least developed and thus present to policymakers what are the "bottlenecks" that hinder the development of the digital entrepreneurial system. After that, the weakest pillar of the digital entrepreneurial system of the Republic of Serbia will be identified, and a comparative analysis of the digital entrepreneurial system of the Republic of Serbia with the member states of the European Union will be presented through the digital aspect of the observed systems.

The purpose of this paper is to diagnose the digital entrepreneurial system of the Republic of Serbia based on its current state, according to the EIDES report, based on which can be concluded what the recommendations for the starting point for its improvement are. As the authors of the EIDES report recommend, the starting points for improving any digital entrepreneurial system are the indicators that slow it down most and hinder its development.

After the introduction, the second part of this paper will present in detail the structure of the digital entrepreneurship system according to the EIDES report, focusing on the digital dimension of the digital entrepreneurship system. Then, in the third part, attention is paid to a detailed description of the methodology, through a description of the research questions, the research sample, and the research instruments. The fourth part focuses on the identification of indicators of the digital entrepreneurial system of the Republic of Serbia that hinder and slow down its development, as well as the presentation of the mathematical calculation which results in the formation of pillars of the digital entrepreneurial system and the identification of the pillar that slows down the development of the digital entrepreneurial system. The fifth part presents a comparative analysis of the digital entrepreneurship system of the Republic of Serbia and digital entrepreneurship systems of EU member states, according to the EIDES report, to conclude whether the Republic of Serbia lags behind in the development of its digital entrepreneurship system. In the final, sixth part, the authors focused on the importance of the analysis as a potential improvement of the regional development of the Republic of Serbia.

2. Materials and Methods

2.1. Research Method

The European Index of Digital Entrepreneurship Systems was first analyzed in a 2018 report [23]. After that, on the same basis, a report for 2019 was prepared [1]. It provides evidence-based support for digital innovation policies and the launch of new reports. When comparing the 2018 and 2019 reports, the 2019 index structure has been adjusted in response to changes in available data. The names of the sub-indexes and pillars of the digital entrepreneurial system have remained the same. However, their variables and the composition of the indicators differ in relation to the 2018 edition.

The authors of the EIDES 2019 report define entrepreneurial ecosystems as follows: "An entrepreneurial ecosystem is a regional community of entrepreneurs, advisors, accelerators and other stakeholders and specialized resources, which support entrepreneurial stand-up, startup and scale-up and entrepreneurial opportunities in the search for digital business models" [1]. This digital entrepreneurial transformation of the economy is very challenging for politics.

The European Index of Digital Entrepreneurship System monitors three groups of conditions in 28 EU member states. However, as mentioned in the introduction, the Republic of Serbia is not a member state and is not covered by the report, which is why the authors chose to focus on the analysis of the entrepreneurial system of the Republic of Serbia.

The European Index of Digital Entrepreneurship Systems monitors three groups of conditions:

- General framework conditions, which describe the context of conducting business in an observed country;
- Systemic framework conditions, which are directly related to entrepreneurial stand-up, startup and scale-up;
- Digital conditions (through the general framework and system framework conditions, i.e., through their digital dimension), which describe the general level of digitalization of an economy because they relate to entrepreneurial activities through the impact on general framework and system framework conditions [1].

The paper's authors focused on analyzing digital conditions, which will be explained in more detail below. The authors of the EIDES report explained precisely how the digital could be distinguished from the non-digital dimension of any entrepreneurial ecosystem, which is analyzed according to their methodology. Namely, when forming the methodology, as previously explained, all indicators are grouped into four large groups: general framework conditions, system framework conditions, digital framework conditions, and digital system conditions. Each of these groups of conditions consists of their associated pillars, as will be explained below. Each pillar is observed in both non-digital and digital dimensions. For example, the pillar of the digital entrepreneurial ecosystem, "culture and informal institutions," is part of the general framework conditions. This pillar can be analyzed in the non-digital dimension and is defined by a specific group of questions from the questionnaire explained below and a particular group of indicators, which correspond to the non-digital dimension of this pillar.

On the other hand, to analyze the pillar "culture and informal institutions" in the digital dimension, it is defined by another group of questions and another group of indicators, which corresponds to the digital dimension, according to the authors of the EIDES report. For the analysis to be relevant, at the very beginning, each pillar of the digital entrepreneurial system must be viewed together with its defining indicators, i.e., it must be considered both in its non-digital and digital dimension, to determine a statistically significant difference between *t*-test and ANOVA observed categories, i.e., pillars of the digital entrepreneurial system. After that prerequisite for further analysis, based on the obtained results, researchers can dedicate themselves to the analysis of the pillars of the digital entrepreneurial system, the analysis of the non-digital dimension, or the analysis of the digital dimension of the digital entrepreneurial system. The paper's authors decided to focus on the digital dimension of the digital entrepreneurial system due to the importance of digitalization in modern business.

The complete concept of The European Index of Digital Entrepreneurship Systems is based on the literature on entrepreneurial ecosystems. The idea of entrepreneurial ecosystems is a new approach, and it can be concluded that there is a relatively weak theoretical basis, which still brings some ambiguities to its concept. The main strength of an entrepreneurial ecosystem lies in its ability to fit many elements of the entrepreneurial context, emphasizing the close relationships and interdependence of these elements to strengthen entrepreneurial mechanisms through different parts of the entrepreneurial ecosystem [24]. The essence of entrepreneurial ecosystems is to take advantage of as many digital benefits as possible, which result from the rapid progress of digital technologies and digital infrastructures, which enable the innovation of business models.

The authors of the EIDES report adopted the concept of an "entrepreneurship system" to distinguish between the unit of analysis at the state level and the unit of analysis at the regional level [22]. The literature on entrepreneurial ecosystems is the most appropriate for understanding digital systems. At the same time, the authors of the EIDES report also emphasize that the analysis at the state level is not relevant.

The EIDES research includes 16 pillars, i.e., 4 pillars derived based on 4 groups within the general framework conditions and 12 pillars that are derived based on system framework conditions, due to differentiation through 3 life phases of the enterprise cycle. The authors of this paper did not focus on a complete analysis of the entrepreneurship system at the level of the state of the Republic of Serbia, based on the EIDES report, but only on its digital conditions.

2.2. Participants

For the purposes of this study, 16 faculties from the territory of the Republic of Serbia were randomly selected to participate in the research. Employees at the faculties were contacted, as well as doctoral students via e-mail, where the questionnaire distribution process was performed according to Dillman's approach [25]. A total of 950 questionnaires were sent, as well as follow-up e-mail reminders, where needed. After 3 months, out of 659 questionnaires received, 300 questionnaires were valid, so that this study includes 300 respondents, of which 136 are women and 154 men, from faculties in the territory of the Republic of Serbia. The sample of 300 respondents from different universities from the territory of the Republic of Serbia seems to be a representative sample. The sample consists of respondents who are doctoral students, researchers, and employees in higher education institutions because they are people who require new scientific information for their doctoral dissertations, which are often related to establishing new companies, creating new or by improving existing business processes, which are vastly more efficient and effective than existing ones, that is, they are related to innovations that can potentially enhance processes in modern business, and thus affect the improvement of the digital entrepreneurial system.

2.3. Data Sample

At the very beginning, a questionnaire and a matrix of all necessary indicators were formed, based on the EIDES report, with the researchers, i.e., the authors adding demographic characteristics and questions about potentially established companies by the respondents [1]. The formation of a questionnaire and matrix with all indicators is essential because it is the only way to adequately arrive at the indicators that slow down the development of the digital entrepreneurial system of the Republic of Serbia.

The questionnaire was sent to employees and doctoral students from the following universities: University of Belgrade, University of Novi Sad, University of Kragujevac, University of Novi Pazar, University of Niš, University of Arts in Belgrade, University of Priština, with temporary headquarters in Kosovska Mitrovica, University of Defense in Belgrade, Alpha BK University, Business Academy University, International University of Novi Pazar, Metropolitan University, Singidunum University, Union University, Educons University, Union-Nikola Tesla University and John Nesbit University.

This research sample was made because employees in higher education institutions and doctoral students are people who very often have innovative ideas that result in starting entrepreneurial companies.

The questionnaire was formed on the LimeSurvey platform and delivered to the respondents electronically. The research was conducted in the period from 1 February 2021 to 1 May 2021.

Respondents were told that the questionnaire was completely anonymous, that they did not need to provide their personal information anywhere, nor that the records of their answers would contain information that could reveal their identity, and that all their answers would be used only for scientific research purposes.

The questionnaire consisted of 47 questions, which were divided into 3 groups: a group referring to demographic data, a group asking about companies potentially established by respondents and their statement whether they are employed or self-employed as well as which alternative they prefer, and a group of questions based on the EIDES questionnaire.

After the answers from the questionnaire were processed, the results from the first group of questions were obtained; that is, the demographic structure of the respondents was formed. The study involved 300 respondents, of which the most significant number belong to the 31 to 40 age group, or 33.67%, while only one respondent belongs to the 18 to 20 age group and only one respondent belongs to the group of respondents aged 71 +, which makes a total of 0.66% of respondents. In the sample, 45.33% were women and 54.67% were men. The most significant number of respondents completed doctoral studies, i.e., 76.67% of respondents. When it comes to the respondents' field of study, the most significant number belong to the technical and technological field, i.e., 53% of respondents. A total of 39% of respondents are from the University of Novi Sad, while 35.67% are from the University of Belgrade.

When it comes to the second group of questions, 3% of respondents said they were "students," 19% of respondents said they were employed in a for-profit organization, 27.67% of respondents were employed in a non-profit organization, 15% of respondents were employed in local or state administration, 4.33% of respondents were self-employed, i.e., own their own company. Furthermore, 72.33% of respondents stated that they prefer to be employed, compared to 27.67% who prefer to be self-employed. A total of 84.33% of respondents had not started a company. In comparison, 7.33% of respondents answered that the companies they own have existed on the market for up to 5 years, 3% of respondents said their companies have existed between 5 and 10 years, and 5.33% of respondents said they founded a company that has been on the market for more than 10 years.

The last, i.e., the third group of questions, was made based on EIDES research and consisted of 36 questions, where, for each statement, respondents were asked to indicate a number that represents the degree of their agreement or disagreement with a statement on a Likert seven-point scale, which was adjusted to the research.

After that, the authors focused on a detailed analysis of the indicators that are the subject of research in this paper, which can be seen below.

2.4. Research Question and Methodology Development

The research is rooted in the basics of the EIDES research, i.e., The European Index of Digital Entrepreneurship Systems, from 2019 [1]. The EIDES report is the basis for the formation of the questionnaire used in the research and for the formation of the matrix, which consists of appropriate indicators and questions from the questionnaire, which provided the basis for evaluation of each pillar in the digital entrepreneurial system of the Republic of Serbia, as well as the basis for pinpointing the advantages and disadvantages of each pillar, which will be explained in more detail below.

The authors decided to focus in detail on the analysis of digital conditions. They observed general framework conditions through non-digital and digital dimensions combined, as well as systemic framework conditions through digital and non-digital dimensions combined to conduct adequate analyses that represent the baseline for further analysis.

General framework conditions consist of 4 pillars:

- Culture and informal institutions;
- Formal institutions, regulations and taxation;
- Market conditions;
- Physical infrastructure.

Systemic framework conditions encompass 4 pillars:

- Human capital;
- Knowledge creation and dissemination;
- Finance;
- Networking and support.

Based on the above, the following research questions were raised:

- RQ1: What are the least developed indicators in the digital dimension of the digital entrepreneurial system of the Republic of Serbia?
- RQ2: What is the least developed category, that is, the pillar of the digital entrepreneurial system of the Republic of Serbia?

Later, when adequate analyses are conducted, and it is determined between which of the above eight categories there is a statistically significant difference, the focus will be placed on the digital dimension of the digital entrepreneurial system of the Republic of Serbia.

3. Results

3.1. Identification of Indicators in the Digital Dimension of Digital Entrepreneurial Systems of the Republic of Serbia

General framework conditions and systemic framework conditions are considered together below, in their non-digital and digital dimensions, to see if and between which pillars of the digital entrepreneurial system of the Republic of Serbia there is a statistically significant difference, which will provide a basis for further detailed analysis.

At the very beginning of the research, the *t*-test assessed whether there is a statistically significant difference between the two groups of indicators: the first group of indicators are general framework conditions with general digital conditions, while the second group of indicators includes systemic framework conditions with systemic digital conditions. Considering that p-value = 0.001, regarding p-value < 0.05, it is concluded that there is a statistically significant difference between these two groups of indicators. The calculation of the mean value supports the conclusion that there is a statistically significant difference between these two groups of indicators, with the additional conclusion that the first group of indicators has a higher mean value compared to the second group, which would mean that the first group of indicators, the General Framework Conditions, in its non-digital and digital dimension, is more developed compared to the second group of indicators, i.e., systemic framework conditions, in its non-digital and digital dimension.

After the *t*-test, ANOVA was used. ANOVA, like the *t*-test, assesses whether there is a statistically significant difference in the mean values of the categories that are being compared, with the *t*-test being able to compare only two categories, while with ANOVA, three or more categories can be compared. With this in mind, ANOVA was used for variables with four and eight categories, which can be seen below.

It was first tested whether there was a statistically significant difference between the categories that make up the general framework conditions, in their non-digital and digital dimension, and the systemic framework conditions, in their non-digital dimension.

General framework conditions consist of the following categories:

- Culture, informal institutions;
- Formal institutions, regulation, taxation;
- Market conditions;
- Physical infrastructure.

Systemic framework conditions consists of the following categories:

- Human capital;
- Knowledge creation and dissemination;
- Finance;
- Networking and support.

As the EIDES report made a detailed matrix of which indicators each category consists of, the paper's authors were guided by it in researching and interpreting the results [1].

Before the ANOVA test, a test of homogeneity of variances was performed first to confirm that the variances within each group were homogeneous. Based on the homogeneity test, it was concluded that the variances are homogeneous because Sig. = 0.376.

Then, the ANOVA test was performed. It was concluded that there is a statistically significant difference between the eight categories mentioned above regarding the indicator's values because Sig. = 0.011, or Sig. < 0.05.

After that, a post hoc test was performed, which served to show between which categories there is a statistically significant difference, and the post hoc test was chosen based on the obtained conclusion that the variances are homogeneous. The LSD post hoc test was selected, the results of which are shown in Table 1.

Table 1. Results of post hoc test.

Category 1	Category 2	*p*-Value
Culture, informal institutions	Market conditions	0.02
Culture, informal institutions	Human capital	0.01
Culture, informal institutions	Knowledge creation and dissemination	0.00
Culture, informal institutions	Finance	0.00
Culture, informal institutions	Networking and support	0.02
Formal institutions, regulations, taxation	Finance	0.02
Physical infrastructure	Finance	0.02

The results of the post hoc test showed exactly between which categories there is a statistically significant difference. Still, the authors wanted to see the value of each category so that they could conclude which category was the most developed or underdeveloped. Therefore, the mean value of each category was calculated. The post hoc test results can be seen below, in Table 2, from which a conclusion can be drawn as to which category is more developed than the others.

Table 2. Results of post hoc test, showing which category is more developed.

Category 1	Category 2	*p*-Value	Category 1 vs. Category 2-Mean
Culture, informal institutions	Market conditions	0.02	1 > 2
Culture, informal institutions	Human capital	0.01	1 > 2
Culture, informal institutions	Knowledge creation and dissemination	0.00	1 > 2
Culture, informal institutions	Finance	0.00	1 > 2
Culture, informal institutions	Networking and support	0.02	1 > 2
Formal institutions, regulations, taxation	Finance	0.02	1 > 2
Physical infrastructure	Finance	0.02	1 > 2

At the beginning of the analysis, general framework conditions and systemic framework conditions were viewed together in their non-digital and digital dimensions, which is adequate for conducting the analyses mentioned above. In the following, when it was determined between which categories there is a statistically significant difference, the paper's authors focused on the analysis of the least developed categories. Between the alternative to observe the least developed indicators in non-digital dimensions and the alternative to observe the least developed indicators within the digital dimension, the authors decided to observe only the least developed indicators within the digital dimension of the digital entrepreneurial system of the Republic of Serbia. Each category consists of appropriate indicators and questions from the questionnaire, formed based on EIDES research. After determining each statistically significant difference between the two observed categories, the less developed category will be identified; its structure will be presented through a table, and the least developed indicator in its digital dimension will be determined.

Observing Culture, informal institutions, and Market conditions, it is concluded that Culture, informal institutions are a category that is more developed than Market conditions. As mentioned above, each category consists of specific indicators. In Table 3, one can see the indicators that make up Market conditions in the digital dimension.

Given that Market conditions is a less developed category, we wanted to see what the "bottleneck" was within this category. Observing the matrix and finding the least developed indicator within the category of Market conditions in the digital dimension, we concluded that the "bottleneck" in the Market conditions category is T-index, with a value of 0.06 in the Republic of Serbia [26]. The T-index is a percentage of the value that estimates the market share of each country in relation to global e-commerce. The higher the T-index, the greater the potential for e-commerce in a given country. In the Republic of Serbia, the T-index is very low compared to the European Union countries, where it is the highest in Germany, with a value of 5.9. The maximum value of the T-index is 100.

Table 3. The components of market conditions in the digital dimension (the table refers to the Republic of Serbia and was made in the likeness of the EIDES report).

Market Conditions in the Digital Dimension of the Digital Entrepreneurial System of the Republic of Serbia		
Indicator	**Indicator Value**	**Maximum Value**
Individuals using the internet for ordering goods or services	34 [27]	100
Enterprises having received orders via computer-mediated networks, % of enterprises	26 [28]	100
Enterprises total turnover from e-commerce	9 [29]	100
Enterprises turnover from web sales	9 [30]	100
T-index	0.06 [26]	100
Pay to advertise on the internet	9 [31]	100

Then, Culture, informal institutions, and Human capital were observed. It was concluded that Culture, informal institutions was a more developed category compared to Human capital. When looking at Human capital, the least developed indicator is the use of the internet to take online courses, with a value of 5 in the Republic of Serbia [32]. In Table 4, one can see the indicators that make up Human capital in the digital dimension. This indicator implies the use of the internet to attend an online course in any subject. The country with the highest value of this indicator in the European Union is Iceland, where the value is 32, and the maximum value is 100.

Table 4. The components of human capital in the digital dimension—the scale-up phase of an enterprise (the table refers to the Republic of Serbia and was made in the likeness of the EIDES report).

Human Capital in the Digital Dimension of the Digital Entrepreneurial System of the Republic of Serbia		
Indicator	**Indicator Value**	**Maximum Value**
Internet use: finding information for goods and services	55 [33]	100
Internet use: doing an online course	5 [32]	100

Then, Culture, informal institutions, and Knowledge creation and dissemination were observed, where the category Culture, informal institutions is more developed than Knowledge creation and dissemination. In Table 5, one can see the indicators that make up Knowledge creation and dissemination in the digital dimension. The least developed digital indicator within the Knowledge creation and dissemination category is software developers, with a value of 0.36 in the Republic of Serbia [34]. Software developers represent the number of software developers per 1000 inhabitants. The highest value of this index is in Germany, with 6.56, while the maximum value is 100.

Table 5. The components knowledge creation and dissemination in the digital dimension—the startup phase of an enterprise (the table refers to the Republic of Serbia and was made in the likeness of the EIDES report).

Knowledge Creation and Dissemination in the Digital Dimension of the Digital Entrepreneurial System of the Republic of Serbia		
Indicator	**Indicator Value**	**Maximum Value**
Employment in high tech and KIBs	4.5 [35]	100
Software developers	0.36 [34]	100

The categories observed next were Culture, informal institutions, and Finance, where it was concluded that Culture, informal institutions is a category that is more developed than the category Finance. In Table 6, one can see the indicators that make up Finance in the digital dimension. It was found that three indicators are underdeveloped within the finance category. As seen below, the Finance category is compared in two more cases with the corresponding categories where there is a statistically significant difference. In this and the following comparison, Formal institutions, regulations, taxation, and Finance will be considered the two least developed indicators of the three least developed indicators within the category of Finance in the digital dimension. We start from the indicators within the digital conditions of Fintech, with a value of 27 in the Republic of Serbia [36]. Fintech represents the number of companies with financial technology per 1,000,000 inhabitants. As a constant maximum value cannot be determined in this indicator, since its limit can be constantly moved, it was observed in relation to the maximum value in the USA, which is 10,605 for financial technology (Fintech).

Table 6. The components of finance in the digital dimension—the scale-up phase of an enterprise (the table refers to the Republic of Serbia and was made in the likeness of the EIDES report).

Finance in the Digital Dimension of the Digital Entrepreneurial System of the Republic of Serbia		
Indicator	**Indicator Value**	**Maximum Value**
Fintech	27 [36]	10,605

Then, the categories Culture, informal institutions, and Networking and supporting were observed. It was concluded that the category Culture, informal institutions, is more developed than Networking and supporting. In Table 7, one can see the indicators that make up Networking and supporting in digital dimension. Observing Networking and supporting within digital conditions, the least developed indicator is the Accelerator amounts, with the value of 0.01 [37]. This is an indicator whose maximum value is 100. The authors of the EIDES research defined it as an indicator obtained when dividing the number of accelerators of the EU member states and GDP per capita.

Table 7. The components of networking and supporting in the digital dimension—the startup phase of an enterprise (the table refers to the Republic of Serbia and was made in the likeness of the EIDES report).

Networking and Supporting in the Digital Dimension of the Digital Entrepreneurial System of the Republic of Serbia		
Indicator	**Indicator Value**	**Maximum Value**
Accelerator number	5 [37]	43
Accelerator amounts	0.01 [37]	100
Meetup Events/Meetup Tech Group Indicator (MTGI)	11 [38]	100
Meetup Members/Meetup Tech Member indicator (MTMI)	1.2 [38]	10

The categories Formal institutions, regulations, taxation, and Finance were observed as well, where it was concluded that Formal institutions, regulations, taxation was more developed than the Finance category. In Table 8, one can see the indicators that make up Finance in the digital dimension. As highlighted in the comparison of the categories Culture, informal institutions and Finance, in this case, the second least developed indicator within the category Finance in digital conditions will be observed. The first indicator is Alternative Financing 3, as defined by the authors of the EIDES research, with a value of 0.05 in the Republic of Serbia [39]. This indicator represents the volume of European business per million inhabitants and is measured in millions of euros. As in this case, it is impossible to determine the maximum fixed value; the maximum value was observed in the countries of the European Union, where France stood out, with 217.78.

Table 8. The components of Finance in the digital dimension—the startup phase of an enterprise (the table refers to the Republic of Serbia and was made in the likeness of the EIDES report).

Finance in the Digital Dimension of the Digital Entrepreneurial System of the Republic of Serbia		
Indicator	**Indicator Value**	**Maximum Value**
Alternative financing 1	0.55 [40]	4.75
Alternative financing 2	0.13 [39]	1.63
Alternative financing 3	0.05 [39]	217.78
Alternative financing 4	1.3 [41]	2.29

Finally, the categories of Psychical infrastructure and Finance were observed, where it was concluded that the category of Psychical infrastructure is more developed than the category of Finance. In Table 9, one can see the indicators that make up Finance in the digital dimension. As previously explained, the third most developed indicator of the three least developed indicators within the category Finance in digital conditions is

Alternative Financing 2, where the authors of the EIDES report defined it as the total alternative financing volume per million inhabitants, where the value of the indicator is 0.13 in the Republic of Serbia [39]. In this case, it is also impossible to determine a constant maximum value, so the maximum value of the countries in the European Union, where France stands out, with the maximum value of 1.63, was taken into account.

Table 9. The components of finance in the digital dimension—the startup phase of an enterprise (the table refers to the Republic of Serbia and was made in the likeness of the EIDES report).

Finance in the Digital Dimension of the Digital Entrepreneurial System of the Republic of Serbia		
Indicator	**Indicator Value**	**Maximum Value**
Alternative financing 1	0.55 [40]	4.75
Alternative financing 2	0.13 [39]	1.63
Alternative financing 3	0.05 [39]	217.78
Alternative financing 4	1.3 [41]	2.29

Indicators in the digital dimension of the digital entrepreneurial system of the Republic of Serbia that can be singled out as "bottlenecks" are those that, to the highest degree, slow down the development of the digital entrepreneurial system of the Republic of Serbia. These are T-index, Usage of the Internet for online courses, Software developers, Fintech, Accelerator amounts, Alternative financing 3, and Alternative financing 2.

The systemic digital conditions of the digital entrepreneurial system of the Republic of Serbia are less developed than the framework digital conditions. Systemic framework conditions represent different resources available to entrepreneurial enterprises in three phases of the enterprise life cycle—stand-up, startup, and scale-up [2]. In Table 10 it can be seen that digital conditions in startup companies are the most underdeveloped. According to the authors of the EIDES report, this is the phase of starting a business and refers to the actual launch of a new entrepreneurial venture and includes early expertise of the business model [1]. Entrepreneurs in the Republic of Serbia, whose companies are in the startup phase, lack Software developers, Alternative financing, and Accelerators the most. They are immediately followed by digital conditions in companies in the scale-up phase. As the authors of the EIDES report state, the scale-up phase encompasses the scale of these entrepreneurial ventures that have uncovered a solid business model, which is scalable. Entrepreneurs in the Republic of Serbia, who have companies in the scale-up phase, have the greatest difficulties using the internet to conduct online courses and fintech.

Table 10. The least developed indicators of the digital entrepreneurial system of the Republic of Serbia in the digital dimension.

	General Framework Conditions			
	Culture, Informal Institutions	**Formal Institutions, Regulations**	**Market Conditions**	**Physical Infrastructure**
			T-index	
	Systemic Framework Conditions			
	Human Capital, Talent	**Knowledge Creations, Dissemination**	**Finance**	**Networking and Support**
Stand-up				
Startup		Software developers	Alternative finance 2 Alternative finance 3	Accelerator amounts
Scale-up	Usage of the Internet for online courses		Fintech	

3.2. Identification of the Pillars of the Digital Entrepreneurial System of the Republic of Serbia

As mentioned above, the paper's authors used the EIDES report to form a survey for the Republic of Serbia. Based on the EIDES report, the values of each pillar of the digital entrepreneurial system in its digital dimension were calculated. The names of the variables and the designations for them are taken from original research.

The difference between general framework conditions and systemic framework conditions is that systemic framework conditions are viewed through three phases of the enterprise life cycle. After each pillar is calculated for each stage of the enterprise life cycle, based on the mean values of the observed pillar in three phases, a value will be calculated that will represent the value of the pillar in all three life phases of the enterprise cycle combined.

General Framework Conditions

DFC_P1 = Culture, informal institutions

$$DFC\_P1 = (DFC\_P1\_I1 + DCP\_P1\_I2 + DFC\_P1\_I3 + DFC\_P1\_I4)/4 \quad (1)$$

$$DFC\_P1 = (0.74 + 0.81 + 0.78 + 0.84)/4$$

Culture, informal institutions = 0.79275

DFC_P2 = Formal Institutions, Regulations and Taxation.

$$DFC\_P2 = (DFC\_P2\_I1 + DCP\_P2\_I2 + DFC\_P2\_I3 + DFC\_P2\_I4 + DFC\_P2\_I5 + DFC\_P2\_I6)/6 \quad (2)$$

$$DFC\_P2 = (0.40 + 0.03 + 0.16 + 0.69 + 0.48 + 0.74)/6$$

Formal Institutions, Regulations and Taxation = 0.41587

DFC_P3 = Market Conditions

$$DFC\_P3 = (DFC\_P3\_I1 + DCP\_P3\_I2 + DFC\_P3\_I3 + DFC\_P3\_I4 + DFC\_P3\_I5 + DFC\_P3\_I6)/6 \quad (3)$$

$$DFC\_P3 = (0.34 + 0.26 + 0.09 + 0.09 + 0.00 + 0.09)/6$$

Market Conditions = 0.1451

DFC_P4 = Physical Infrastructure

$$DFC\_P4 = (DFC\_P4\_I1 + DCP\_P4\_I2 + DFC\_P4\_I3 + DFC\_P4\_I4 + DFC\_P4\_I5 + DFC\_P4\_I6 + DFC\_P4\_I7)/7 \quad (4)$$

$$DFC\_P4 = (0.29 + 0.19 + 0.04 + 0.06 + 0.34 + 1.00 + 0.93)/7$$

Physical Infrastructure = 0.406531

Systemic Framework Conditions

Stand-up

S1_SDC_P1 = Human capital

$$S1\_SDC\_P1 = (S1\_SDC\_P1\_I1 + S1\_SDC\_P1\_I2 + S1\_SDC\_P1\_I3)/3 \quad (5)$$

$$S1\_SDC\_P1 = (0.09 + 0.50 + 0.46)/3$$

Human capital = 0.350317

S1_SDC_P2 = Knowledge Creation and Dissemination

$$S1\_SDC\_P2 = (S1\_SDC\_P2\_I1 + S1\_SDC\_P2\_I2)/2 \quad (6)$$

$$S1\_SDC\_P2 = (1.00 + 0.27)/2$$

Knowledge Creation and Dissemination = 0.63495

S1_SDC_P3 = Finance

$$S1\_SDC\_P3 = (S1\_SDC\_P3\_I1 + S1\_SDC\_P3\_I2)/2 \quad (7)$$

$$S1\_SDC\_P3 = (0.17 + 0.18)/2$$

Finance = 0.17333

S1_SDC_P4 = Networking and Support

$$S1\_SDC\_P4 = (S1\_SDC\_P4\_I1 + S1\_SDC\_P4\_I2 + S1\_SDC\_P4\_I3 + S1\_SDC\_P4\_I4)/4 \quad (8)$$

$$S1\_SDC\_P4 = (0.49 + 0.55 + 0.11)/3$$

Networking and Support = 0.419167
Startup
S2_SDC_P1 = Human capital

$$S2\_SDC\_P1 = (S2\_SDC\_P1\_I1)/1 \quad (9)$$

$$S2\_SDC\_P1 = (0.21)/1$$

Human capital = 0.21
S2_SDC_P2 = Knowledge Creation and Dissemination

$$S2\_SDC\_P2 = (S2\_SDC\_P2\_I1 + S2\_SDC\_P2\_I2)/2 \quad (10)$$

$$S2\_SDC\_P2 = (0.05 + 0.00)/2$$

Knowledge Creation and Dissemination = 0.0243
S2_SDC_P3 = Finance

$$S2\_SDC\_P3 = (S2\_SDC\_P3\_I1 + S2\_SDC\_P3\_I2 + S2\_SDC\_P3\_I3 + S2\_SDC\_P3\_I4)/4 \quad (11)$$

$$S2\_SDC\_P3 = (0.12 + 0.08 + 0.00 + 0.57)/3$$

Finance = 0.190865
S2_SDC_P4 = Networking and Support

$$S2\_SDC\_P4 = (S2\_SDC\_P4\_I1 + S2\_SDC\_P4\_I2 + S2\_SDC\_P4\_I3 + S2\_SDC\_P4\_I4)/4 \quad (12)$$

$$S2\_SDC\_P4 = (0.12 + 0.00 + 0.11 + 0.12)/4$$

Networking and Support = 0.086595
Scale-up
S3_SDC_P1 = Human capital

$$S3\_SDC\_P1 = (S3\_SDC\_P1\_I1 + S3\_SDC\_P1\_I2)/2 \quad (13)$$

$$S3\_SDC\_P1 = (0.55 + 0.05)/2$$

Human capital = 0.3
S3_SDC_P2 = Knowledge Creation and Dissemination

$$S3\_SDC\_P2 = (S3\_SDC\_P2\_I1 + S3\_SDC\_P2\_I2)/2 \quad (14)$$

$$S3\_SDC\_P2 = (0.29 + 0.29)/2$$

Knowledge Creation and Dissemination = 0.29
S3_SDC_P3 = Finance

$$S3\_SDC\_P3 = (S3\_SDC\_P3\_I1)/1 \quad (15)$$

$$S3\_SDC\_P3 = (0.00)/1$$

Finance = 0.00
S3_SDC_P4 = Networking and Support

$$S3\_SDC\_P4 = (S3\_SDC\_P4\_I1 + S3\_SDC\_P4\_I2)/2 \quad (16)$$

$$S3\_SDC\_P4 = (0.13 + 0.29)/2$$

Networking and Support = 0.21
SDC_P1 = Human capital, using (5), (9) and (13)

$$SDC\_P1 = (S1\_SDC\_P1 + S2\_SDC\_P1 + S3\_SDC\_P1)/3$$

$$SDC\_P1 = (0.350317 + 0.21 + 0.3)/3$$

Human capital = 0.286105
SDC_P2 = Knowledge Creation and Dissemination, using (6), (10) and (14)

$$SDC\_P2 = (S1\_SDC\_P2 + S2\_SDC\_P2 + S3\_SDC\_P2)/3$$

$$SDC\_P2 = (0.63495 + 0.0243 + 0.29)/3$$

Knowledge Creation and Dissemination = 0.316416
SDC_P3 = Finance, using (7), (11) and (15)

$$SDC\_P3 = (S1\_SDC\_P3 + S2\_SDC\_P3 + S3\_SDC\_P3)/3$$

$$S3\_SDC\_P3 = (0.17333 + 0.190865 + 0.00)/3$$

Finance = 0.00
SDC_P4 = Networking and Support, using (8), (12) and (16)

$$SDC\_P4 = (S1\_SDC\_P4\_ + S2\_SDC\_P4 + S3\_SDC\_P4)/3$$

$$S3\_SDC\_P4 = (0.419167 + 0.086595 + 0.21)/3$$

Networking and Support = 0.21
Based on the mathematical calculation above and Table 11, it is concluded that the Finance pillar of the digital entrepreneurial system is the least developed in its digital dimension, with a value of 0.12224.

Table 11. The pillars of the digital entrepreneurial system of the Republic of Serbia, in their digital dimension, with the identification of the least developed pillar.

Digital Pillars of the Digital Entrepreneurial System of the Republic of Serbia	
Pillar	**Pillar Value**
Culture and Informal Institutions	0.79275
Formal Institutions, Regulations and Taxation	0.41586
Market Conditions	0.14510
Physical Infrastructure	0.40653
Human Capital	0.28610
Knowledge Creation and Dissemination	0.31641
Finance	0.12224
Networking and Support	0.23858

All values in the table are normalized, where the lowest value is 0, while the highest is 1, making it easier for the reader to understand how weak or strong each pillar is.

4. Discussion

The Republic of Serbia, according to the Statistical Office of the Republic of Serbia, and according to the 2019 census, has 6,945,235 inhabitants [42]. According to the World Bank, GDP per capita for 2020 is USD 7673, or 755,335 Serbian dinars [43].

As shown in Figure 1 and Table 11, the following classification of the pillars of the digital entrepreneurial system of the Republic of Serbia, observed in its digital dimension, from the least developed to the most developed, can be made:

- Finance;
- Market conditions;
- Networking and support;
- Human capital;
- Knowledge creation and dissemination;
- Physical infrastructure;
- Formal institutions;
- Culture and informal institutions.

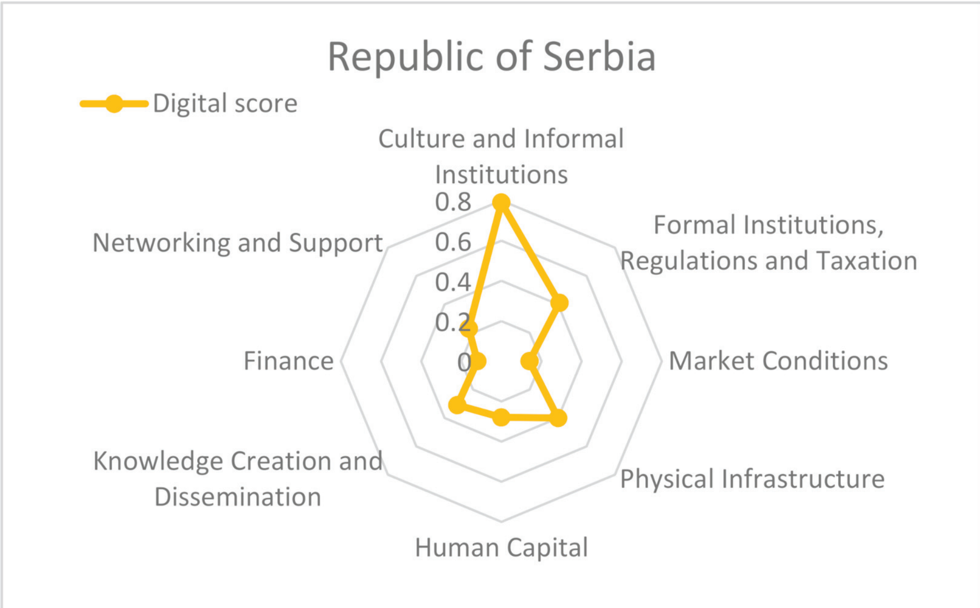

Figure 1. Values of pillars in the digital entrepreneurial system of the Republic of Serbia, in their digital dimension.

Evaluation of the digital entrepreneurial system of the Republic of Serbia in its digital dimension is calculated as a mean value and is 34, while the highest score is 100.

Finances affect the availability of different forms of financing for new ventures. As it could be concluded from the previous part where the least developed indicators within the category of Finance were analyzed, entrepreneurs in the Republic of Serbia have the biggest problem with alternative financing and fintech. There are not enough available alternative sources of funding in the Republic of Serbia, which would help entrepreneurs in the startup and scale-up phase of their entrepreneurial business ventures to overcome the crises that await them. On the other hand, Fintech, as a symbiosis of technologies and customer services for the financial market, is also a "bottleneck" of the digital entrepreneurial system of the Republic of Serbia. It can be concluded that Finance is the primary field in which policymakers should work to enable further development of the digital entrepreneurial system of the Republic of Serbia.

The most developed pillar of the digital entrepreneurial system of the Republic of Serbia, in its digital dimension, is Culture, informal institutions, that is, a group of indicators that affect the individual attitude towards entrepreneurship as a career choice.

Comparing the digital dimension of digital entrepreneurial systems of the EU member states, from the EIDES report, it is concluded that the Republic of Serbia is lagging behind the EU member states. In Table 12, you can see where the Republic of Serbia is compared to the European Union member states.

Table 12. Comparison of the Republic of Serbia and EU member states, based on the digital entrepreneurial system.

Digital Entrepreneurial System, Viewed in the Digital Dimension	
Country	**Score**
Republic of Serbia	34.0
Greece	47.4
Bulgaria	47.8
Romania	50.5
Croatia	54.4
Italy	55.6
Portugal	56.5
Latvia	57.2
Poland	57.2
Cyprus	58.2
Slovakia	58.8
Hungary	59.9
Slovenia	62.1
Spain	65.8
Czech Republic	66.1
Lithuania	66.6
France	67.1
Austria	70.1
Malta	70.8
Ireland	73.7
Belgium	73.8
Estonia	74.4
Germany	78.7
Finland	81.3
Luxembourg	81.8
Denmark	83.4
United Kingdom	85.1
Sweden	85.2
The Netherlands	86.3

The authors of this research focused on the digital dimension of the digital entrepreneurial system of the Republic of Serbia. Still, it is recommended that the entrepreneurial system policy consider both stand-up, startup and scale-up comprehensively and consider the dynamics of the ecosystem as a whole.

Once the "bottlenecks" of the digital entrepreneurial system have been identified, they must actively involve all ecosystem stakeholders for entrepreneurial ecosystem policies to be successful. Entrepreneurial ecosystem policies require a long-term approach.

The policy approach to the entrepreneurial ecosystem is likely to pose significant challenges to policy-making and implementation institutions as they are in many cases shaped in the traditional way of decision-making. One of the proposals for overcoming this problem is to form partnerships with regional organizations, which have the authority and are sufficiently committed to undertaking long-term processes to help improve entrepreneurial ecosystems.

The authors propose several solutions for the successful implementation of the improvement of entrepreneurial ecosystems:

- To combine regional and national approaches;
- That the structures and processes of the entrepreneurial ecosystem differ at different levels of policy implementation;
- To foster learning and exchange of experiences across the region;
- To approach the entrepreneurial ecosystem in such a way as to understand how entrepreneurial ecosystems work [44].

As mentioned at the outset, entrepreneurial ecosystems are driven by the process of digitalization and close coordination between the entrepreneurial ecosystem and digitalization policies is needed.

5. Conclusions

Entrepreneurial ecosystems can be viewed as systems whose purpose is to enable and facilitate the allocation of resources for productive purposes. They are made possible by the ubiquitous trend of digitalization, which is advancing day by day and constantly opens up opportunities to re-examine an organization and whether it creates new values in the digital economy through innovations in its business model. This is precisely the feature that makes entrepreneurial ecosystems a key driver of progress towards the digital economy [4].

The characteristics of the entrepreneurial ecosystem make the entrepreneurial ecosystem a vital policy challenge because the entrepreneurial ecosystem should be viewed as a systemic process, the dynamics of which are not easy to reduce to actions at the enterprise level, which can be easily concluded from the presented research.

Entrepreneurial ecosystems can be understood as a process that is mainly observed at the regional level. They consist of regional stakeholders and resources that specialize in facilitating the assessment of stand-up, startup, and scale-up of new ventures. Since the authors of the EIDES report took into account that this regional dimension is unavailable due to lack of supplements, they formed an EIDES score based on pillars grouped into general and systemic framework conditions governing entrepreneurial ecosystem dynamics. As general conditions represent conditions at the national level that are more or less similarly applied to all regional clusters, they are better suited for resolution at the state level. Systemic conditions tend to show more regional variations because they represent the characteristics of regional communities. That is why the EIDES index measures both types of conditions. It uses data at the national level for general conditions and actions at the regional level to solve systemic framework conditions, which was an excellent starting point for researching the digital entrepreneurial system of Serbia.

Author Contributions: Conceptualization, J.R., Đ.Ć. and B.D.; methodology, J.Ć.; software, D.S.; validation, J.R., D.S. and B.D.; formal analysis, Đ.Ć.; investigation, J.Ć.; resources, B.D. and J.Ć.; data curation, D.S.; writing—original draft preparation, J.R.; writing—review and editing, Đ.Ć.; visualization, B.D.; supervision, D.S.; project administration, J.R. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data is contained within the article.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Autio, E.; Szerb, L.; Komlósi, E.; Tiszberger, M. *EIDES 2019—The European Index of Digital Entrepreneurship Systems*; EUR 29892 EN; Publications Office of the European Union: Luxembourg, 2019; ISBN 978-92-76-12269-2. [CrossRef]
2. Autio, E.; Rannikko, H. Entrepreneurial Dynamic of the Digital Economy and Finland's International Competitiveness. In *Publications of the Government's Analysis, Assessment and Research Activities*; Prime Minister's Office, Ed.; Prime Minister's Office: Helsinki, Finland, 2017.
3. Carlsson, B. The Digital Economy: What is new and what is not? *Struct. Chang. Econ. Dyn.* **2004**, *15*, 245–264. [CrossRef]
4. Autio, E.; Nambisan, S.; Thomas, L.; Wright, M. Digital affordances, spatial affordances, and the genesis of entrepreneurial ecosystems. *Strateg. Entrep. J.* **2017**, *12*, 72–95. [CrossRef]
5. Hair, N.; Wetsch, L.; Hull, C.; Perotti, V.; Hung, Y.-T. Market orientation digital entrepreneurship: Advantages and challenges web 2.0 networked world. *Int. J. Innov. Technol. Manag.* **2012**, *9*, 1–17. [CrossRef]
6. Guthrie, C. The digital factory: A hands-on learning project digital entrepreneurship. *J. Entrep. Educ.* **2014**, *17*, 115–133.

7. Le Dinh, T.; Vu, M.C.; Ayayi, A. Towards a living lab for promoting the digital entrepreneurship process. *Int. J. Entrep.* **2018**, *22*, 1–17.
8. Davidson, E.; Vaast, E. Digital entrepreneurship and its Sociomaterial enactment. In Proceedings of the 43rd Hawaii International Conference on System Sciences, Koloa, HI, USA, 5–8 January 2010.
9. Millman, C.; Wong, W.-C.; Li, Z.; Matlay, H. Educating students for e-entrepreneurship in the UK, the USA, and China. *Ind. High. Educ.* **2009**, *23*, 243–252. [CrossRef]
10. Autio, E.; Cao, S.; Temsiripoj, W.; Guntasopatr, P.; Kaensup, P.; Chumjit, S. *Enhancing Thailand's Ecosystems for Entrepreneurship*; King Mongkut University of Technology: Bangkok, Thailand, 2018.
11. Stam, E. The Dutch Entrepreneurial Ecosystem. SSRN. 2014. Available online: https://ssrn.com/abstract=2473475 (accessed on 31 March 2021).
12. Szerb, L.; Acs, Z.; Autio, E.; Ortega-Argiles, E.; Komplosi, E. *REDI: The Regional Entrepreneurship and Development Index—Measuring Regional Entrepreneurship*, 1st ed.; European Commission, Directorate-General for Regional and Urban Policy, Publications Office of the European Union: Luxembourg, 2014.
13. Audretsch, D.; Link, A. Valuing an entrepreneurial enterprise. In *Concise Guide to Entrepreneurship, Technology, and Innovation*, 1st ed.; Audretsch, D., Hayter, S., Link, N.A., Eds.; Edward Elgar: Chelterham, PA, USA, 2015.
14. Stam, E.; Nooteboom, B. Entrepreneurship, innovation and institutions. In *Handbook of Research on Innovation and Entrepreneurship*, 1st ed.; Audretsch, D., Falck, O., Hevlich, S., Eds.; Edward Elgar: Cheltenham, PA, USA, 2011; pp. 421–436.
15. Foster, G.; Shimizu, C.; Ciesinski, S.; Davila, A.; Hassan, S.; Jia, N.; Morris, R. Entrepreneurial ecosystems around the globe and company growth dynamics. In *World Economic Forum*; World Economic Forum: Geneva, Switzerland, 2013; Volume 11, pp. 1–36.
16. Audretsch, D.; Lehmann, E. *The Seven Secrets of Germany: Economic Resilience in an Era of Global Turbulence*, 1st ed.; Oxford University Press: Oxford, UK, 2016.
17. Szerb, L.; Acs, Z.; Komlosi, E.; Ortega-Argiles, R. Discussion Paper—Measuring Entrepreneurial Ecosystems: The Regional Entrepreneurship and Development *Index (REDI)*. Discussion Paper Number: CFE-2015-02. 2015. Available online: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.724.3755&rep=rep1&type=pdf (accessed on 31 December 2020).
18. Spigel, B. The relational organization of entrepreneurial ecosystem. *Entrep. Theory Pract.* **2017**, *41*, 49–72. [CrossRef]
19. Matthews, C.H.; Brueggemann, R.F. Innovation entrepreneurship a competency framework. In *Innovation and Entrepreneurship: A Competency Framework*; Routledge: New York, NY, USA, 2015.
20. Acs, Z.; Szerb, L.; Lloyd, A. *Global Entrepreneurship Index Powered by GEDI*; CreateSpace Independent Publishing Platform: Washington, DC, USA, 2018.
21. Elia, G.; Margherita, A.; Passiante, G. Digital entrepreneurship ecosystem: How digital technologies and collective intelligence are reshaping the entrepreneurial process. *Technol. Forecast. Soc. Chang.* **2020**, *150*, 119791. [CrossRef]
22. Acs, Z.J.; Autio, E.; Szerb, L. National Systems of Entrepreneurship: Measurement Issues and Policy Implications. *Res. Policy* **2014**, *43*, 476–494. [CrossRef]
23. Autio, E.; Szerb, L.; Komlósi, E.; Tiszberger, M. *The European Index of Digital Entrepreneurship Systems*; EUR 29309 EN; Publications Office of the European Union: Luxembourg, 2018; p. JRC112439. ISBN 978-92-79-91303-7. [CrossRef]
24. Autio, E.; Wright, M. Entrepreneurial innovation: The importance of context. *Res. Policy* **2014**, *43*, 1097–1108. [CrossRef]
25. Dillman, D.A. *Mail and Internet Surveys: The Tailored Design Method-2007 Update with New Internet, Visual, and Mixed-Mote Guide*; John Wiley & Sons: New York, NY, USA, 2011.
26. Translated. The Top Countries for Online Sales. Available online: https://translated.com/languages-that-matter (accessed on 5 April 2021).
27. Eurostat. Individuals Using the Internet for Ordering Goods or Services. Available online: https://ec.europa.eu/eurostat/databrowser/view/tin00096/default/table?lang=en (accessed on 5 April 2021).
28. Eurostat. E-Commerce Sales. Available online: https://appsso.eurostat.ec.europa.eu/nui/show.do?dataset=isoc_ec_eseln2&lang=en (accessed on 5 April 2021).
29. Eurostat. Share of Enterprises' Turnover on E-Commerce—%. Available online: https://ec.europa.eu/eurostat/databrowser/view/tin00110/default/table?lang=en (accessed on 5 April 2021).
30. Eurostat. Enterprises Turnover from Web Sales. Available online: https://appsso.eurostat.ec.europa.eu/nui/setupModifyTableLayout.do (accessed on 5 April 2021).
31. Eurostat. Value of E-Commerce Sales. Available online: https://appsso.eurostat.ec.europa.eu/nui/submitViewTableAction.do (accessed on 5 April 2021).
32. Eurostat. Individuals Using the Internet for Doing an Online Course. Available online: https://ec.europa.eu/eurostat/databrowser/view/tin00103/default/table?lang=en (accessed on 5 April 2021).
33. Eurostat. Individuals Using the Internet for Finding Information about Goods and Services. Available online: https://ec.europa.eu/eurostat/databrowser/view/tin00095/default/table?lang=en (accessed on 5 April 2021).
34. Stack Overflow. Developer Survey Results. Available online: https://insights.stackoverflow.com/survey/2018 (accessed on 5 April 2021).
35. Eurostat. Employment in High Tech and KIBs. Available online: http://ec.europa.eu/eurostat/tgm/table.do?tab=table&init=1&laguage=en&pcode=tsc00011&plugin=1 (accessed on 5 April 2021).

36. Deal Room. Fintech. Available online: https://app.dealroom.co/companies/f/_10/anyof_2017)/industries/anyof_fintech/locations/anyof_Europe%20(20_Serbia (accessed on 5 April 2021).
37. Startup Blink. Accelerator Amounts. Available online: https://www.startupblink.com/accelerators/serbia (accessed on 5 April 2021).
38. Meet Up. Dive In! There Are So Many Things to Do on Meetup. Available online: https://www.meetup.com (accessed on 5 April 2021).
39. Ziegler, T.; Shneor, R.; Garvey, K.; Wenzlaff, K.; Yerolemou, N.; Rui, H.; Zhang, B. Expanding Horizons: The 3rd European Alternative Finance Industry Report. *SSRN Electron. J* **2018**. Available online: https://ssrn.com/abstract=3106911 (accessed on 31 January 2021). [CrossRef]
40. Zhang, B.; Wardrop, R.; Ziegler, T.; Lui, A.; Burton, J.; James, A.; Garvey, K. *Sustaining Momentum: The 2nd European Alternative Finance Industry Report*; Cambridge Centre for Alternative Finance: Cambridge, UK, 2016; Available online: https://www.jbs.cam.ac.uk/fileadmin/user_upload/research/centres/alternative-finance/downloads/2016-european-alternative-finance-report-sustaining-momentum.pdf (accessed on 5 April 2021).
41. Statista. Alternative Financing—Serbia. Available online: https://www.statista.com/outlook/dmo/fintech/alternative-financing/serbia (accessed on 5 April 2021).
42. Republic Statistical Office of the Republic of Serbia. Procene Stanovništva. Available online: https://www.stat.gov.rs/sr-Latn/oblasti/stanovnistvo/procene-stanovnistva (accessed on 5 April 2021).
43. World Bank. The World Bank in Serbia. Available online: https://www.worldbank.org/en/country/serbia/overview (accessed on 5 April 2021).
44. Autio, E.; Rannikko, H. Retaining winners: Can policy boost high-growth entrepreneurship? *Res. Policy* **2016**, *45*, 42–55. [CrossRef]

mathematics

MDPI

Article

Unsustainability Risk of Bid Bonds in Public Tenders

Jacopo Giacomelli [1,2,*] and Luca Passalacqua [2]

1 SACE S.p.A-Piazza Poli 42, 00187 Rome, Italy
2 Department of Statistics, Sapienza University of Rome,Viale Regina Elena 295, 00161 Rome, Italy; luca.passalacqua@uniroma1.it
* Correspondence: j.giacomelli@sace.it

Abstract: Public works contracts are commonly priced and awarded through a tender process. Each bidder joining the tender must underwrite a bid bond that guarantees their fitness as contractors in case of a win. The winning contractor also needs to underwrite a performance bond before entering the contract to protect the procuring entity against the performance risk arising during the execution phase. This study addresses the case when sureties refuse to issue the performance bond, despite having issued a bid bond to the same subject. A creditworthiness variation of the contractor during the tender or an excessive discount of the contract's price may lead to this outcome. In that case, all the subjects involved are damaged. The surety who issued the bid bond has to indemnify the procuring entity. The contract award is nullified, which is financially harmful to both the contractor and the procuring entity. We show that sureties adopting a forward-looking risk appetite framework may prevent the demand for unsustainable performance bonds instead of addressing it by rejecting the bidders' requests. The Solvency II regulatory framework, the Italian bidding law, and actual historical data available from the Italian construction sector are considered to specify a simplified model. The probability of unsustainable tender outcomes is numerically estimated by the model, together with the mitigating impact of a surety's proper strategy.

Keywords: bid bond; suretyship; risk management; decision under uncertainty; Solvency II

1. Introduction

Sustainability is a complex and evolving concept that may include, inter alia, economic and financial considerations, environmental and social impacts, as well as political and legal aspects [1]. In public works, sustainability must be considered in its broadest meaning since all the aspects are relevant to the public entity that requires the execution of a project and the citizens who benefit from its fulfillment. However, also in this context, economic sustainability remains necessary to enable all the other possible forms of sustainability. Unfair pricing of the project implies a waste of public resources (when too high) or may result in a poor or even missed execution (when too low). Both cases have a negative impact, at least from a social perspective, but possibly also environmentally, depending on the specific situation.

Typically, the cost of a public construction project is determined by a tender promoted by the procuring entity. Nowadays, each country disciplines bid mechanisms underlying public tenders by a complex regulatory framework that guarantees fairness among participants and financial protection to the procuring entities. The leading economies share the main features of respective public bid laws (see, e.g., [2–5]). In particular, a system of guarantees is usually mandatory and involves insurance companies and financial institutions as sureties [6–8]. Each participant to a public tender must underwrite a bid bond to take part in the tender. The bid bond guarantees that the contract winner will satisfy all the requirements needed to become the contractor, including acquiring a performance bond. The values of the contract and the related performance bond are subjected to stochastic variations during the bidding process due to the tender rules. Hence, when the bid bond is issued, the surety has to consider the riskiness of the guaranteed participant concerning these variations and the sustainability of issuing a subsequent performance bond in case the bidder wins the contract.

There is a negative dependency between the final value of the contract and the notional value of the performance bond. This is because the public procuring entity wants to be protected against the performance risk of the winning contractor. Intuitively, the lower the final performance cost is, the higher the probability of poor performance or other breaches of contract is. Thus, the need for financial protection of the public entity increases accordingly. The final notional value of the performance bond may be too high for the risk appetite of the surety who has issued the bid bond of the winning bidder, significantly if the bidder's creditworthiness has worsened during the tender process. If no other surety is available to issue the performance bond, the bid bond generates a claim. The surety who has issued it has to indemnify the procuring entity on behalf of the contractor, who cannot be awarded the contract.

During the last two decades, the increasing need for developing and improving public infrastructures in many countries has renewed the research interest on various financial, economic, and legal topics related to the construction industry. In particular, the attention to risk management tools and techniques in public works and large private projects has considerably increased, leading to an intense research activity [9–12]. In this context, surety bonds have been investigated mainly empirically, with particular reference to performance bonds and the benefits they produce in terms of risk mitigation. In these years, performance bonds have also been investigated with regards to their legal sustainability, depending on the specific regulatory framework of a considered country [13–15]. For example, in countries where these instruments have been introduced recently, it is worth considering the moral hazard of beneficiaries who abuse their right to call on the surety guarantees. The relevance of the problem has been assessed, and possible improvements to specific national laws are presented in [13,14]. Surety bonds have been investigated from an actuarial perspective as well, addressing both the problems of pricing them and measuring their mitigation effect on the underlying performance risk [16,17]. Results obtained in [16] imply that surety companies can help to mitigate the problem of contractors going bankrupt by their ability to perform a preliminary screening. Further, in [17] it is shown that contractors with a better standing are more likely to win the tender if sureties apply a risk-adjusted price to the performance bond. However, to date, the literature has focused only on investigating the performance risk and the likelihood that the winning contractor defaults during the execution of the public works.

This work investigates the case that the tender process may lead to an unsustainable outcome (i.e., the performance bond is not issued, and the tender process has to be reopened). This is relevant since the inefficiency of the tender process implies costs for all the subjects involved: the bidders, the sureties, and the public entity. Italian bid law [18,19] is considered to specify the tender mechanisms (e.g., the functional form that links the contract pricing and the notional value of the performance bond). A risk appetite framework is proposed to model the behavior of the sureties who support the bidders based on the Solvency II regulatory framework [20–22]. To the best of our knowledge, this is the first study that addresses this specific topic.

The work is organized as follows. Section 2 introduces the business and legal context of the investigated problem. The main features of the suretyship insurance business are reported, with a focus on the bid bond and the performance bond insurance products. Further, the Italian bidding law for public works is described. Section 3 models the sustainability of the tender outcome both for the bidders and the sureties. The Solvency II Standard Formula elements needed to design the surety's risk appetite framework are introduced, and the bidder's behavior is modeled considering their appetite for a minimum profit. Section 4 addresses the investigated problem. After introducing the distributional assumptions needed, we measure the probability of inefficient outcomes of public tenders through numerical simulations. The main results are summarized in Section 5.

2. Elements of Suretyship Insurance and the Italian Public Tenders

Suretyship policies provide a guarantee of performance and principles of various objectives and duties. They are commonly required to secure the obligations of the principal debtor (generally known as the *principal*) against the beneficiary. In the Solvency II framework, suretyship insurance is classified in *Solvency II Line of Business* 9 (also known as S2LoB 9), together with credit insurance. However, unlike in credit insurance, the suretyship insurer (also known as the *surety*) has a direct relation with the source of risk. The principal usually underwrites a surety policy because this is a requirement to engage the beneficiary in business.

Risks underlying surety policies can be very diverse from each other, ranging from performance risk in an engineering contract to moral hazard/operational risks in claiming a VAT credit to be refunded. Surety bonds fall under two categories [6]:

- Contract bonds, intended to guarantee the performance of contractual obligations, mainly in the areas of public works and private construction projects;
- Commercial bonds, intended to secure the performance of legal or regulatory obligations.

Without claim to completeness, examples of products belonging to each category are listed in Table 1.

Table 1. A brief description of the main suretyship products.

Contract Bonds	
Bid bond	Guarantees that a contractor has submitted a bid in good faith and intends to enter the contract in case of award
Performance bond	Offers protection from the case that a contractor fails to fulfill the terms of the contract
Advance payment bond	Guarantees that the contractor will be able to repay the procuring entity any funds received in advance
Payment bond	Protects the credit of workers, subcontractors, and suppliers against the contractor
Maintenance bond	Guarantees against defective workmanship or materials
Commercial Bonds	
Customs bond	Assures customs authorities that an importer will pay the import duties required
Tax bond	Ensures the proper declaration and timely payment of taxes
License/permit bond	Guarantees the obligor's compliance with laws
Court/fidelity bond	Guarantees the performance of fiduciaries' duties and their compliance with court orders

This work is focused on two typical products in suretyship insurance among the ones listed above: the *bid bond* and the *performance bond*. The life cycle of these two products—depicted in Figures 1 and 2—can be summarized as follows.

A procuring entity requires a generic "performance", such as constructing infrastructures or supplying specific goods or services. Hence, the entity mentioned above uses a bidding process to select the best contractor for the assignment. Each contractor interested in submitting the bid has to underwrite a *bid bond* that guarantees the procuring entity against the case that the awarded contractor is not able to take charge of the required performance. Indeed, the contractor could go bankrupt during the bidding process, or some requirement necessary to fulfill the obligation could not be met (e.g., legal authorizations needed to perform the underlying task). In this case, the bidding process has to be reopened, and the insurer indemnifies the procuring entity.

In case the winning contractor satisfies all the requirements, a *performance bond* is still needed to close the bidding process. It is worth noticing that the insurer who has issued the bid bond may refuse to underwrite the performance bond. However, suppose the contractor cannot find another insurer available to underwrite the required performance bond. In that case, the tender is reopened, and thus the bid bond issuer has to indemnify the procuring entity. This mechanism implies that insurers who issue a bid bond share the subsequent performance risk with the procuring entity to some extent.

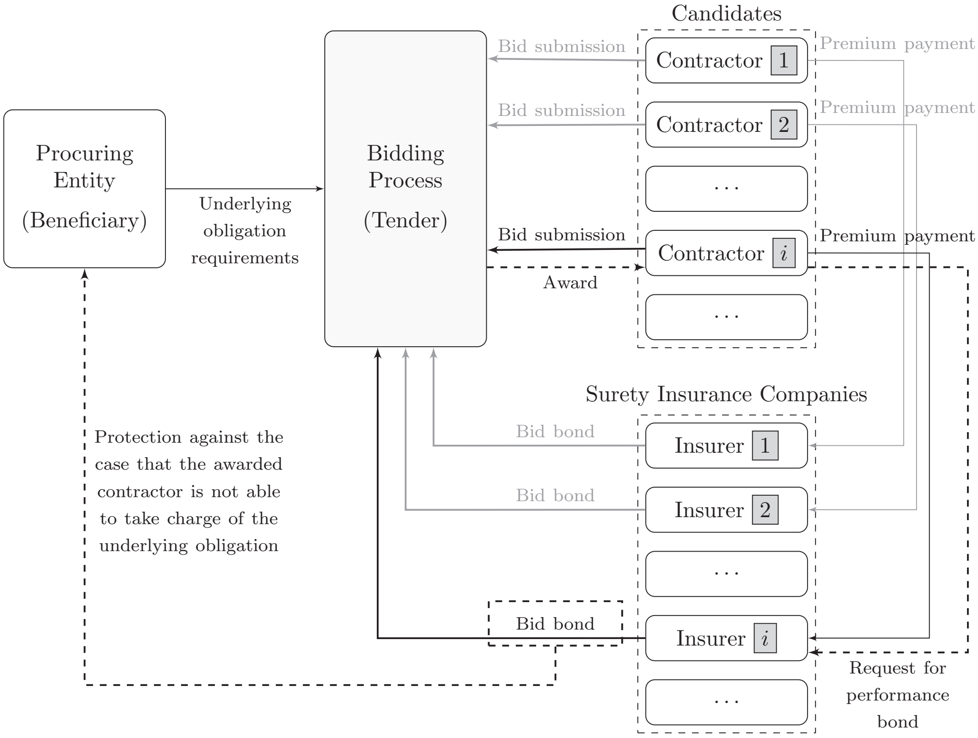

Figure 1. Schematics of the *bid bond*. The bidding process is represented, in which the bid bond protects the beneficiary and the following performance bond is not issued yet.

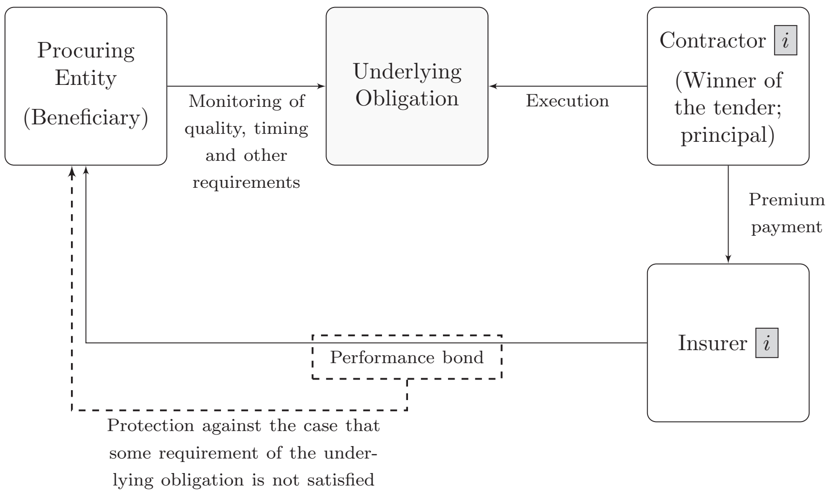

Figure 2. Schematics of the *performance bond*. The execution phase is represented, in which the performance bond protects the beneficiary.

The phase after the bidding process is the *execution phase* when the obligation has to be fulfilled by the winner of the tender. The performance bond guarantees the beneficiary against the risk that the principal cannot satisfy the obligation's timing or any other requirement. If the performance does not meet all the requirements declared in the bidding process, the insurer indemnifies the procuring entity.

Depending on the considered regulatory framework, the procuring entity can increase the duration or modify other features of the obligation during the execution phase. This is usually the case when the procuring entity is a public institution. When the risk underlying the performance bond is modified, the insurer may require the payment of a premium supplement but must accept to guarantee the beneficiary against the risk extension.

In the case of a claim, the *subrogation* phase takes place. Namely, the subrogation can be thought of as the set of rules that defines the insurer's role when the execution phase is interrupted by a violation of the underlying obligation. Different regulatory frameworks define this concept in different ways.

Bid and Performance Bonds: The Italian Case

According to the Italian bidding law [18], the exposure (in this study, "exposure" refers to the maximum claim size that a considered insurance contract covers without considering any recoverable) N_B generated by a bid bond is typically fixed to a $\alpha_B = 2\%$ fraction of the underlying obligation notional value V_0, established by the procuring entity at the beginning $t = t_0$ of the bidding process. However, depending on the risk profile of the obligation, the procuring entity may choose a different α_B value in the interval $[1\%, 4\%]$.

The performance bond exposure N_P likewise is represented as a fraction α_P of the obligation notional values V_1, re-estabilished at the time $t = t_1$ when the tender ends. Since each candidate contractor offers to take charge of the obligation at a cost that is lower than the auction base V_0, it holds $V_1 < V_0$ by construction. However, the Italian law forbids to choose α_P and provides a mechanism to protect the beneficiary from the risk arising when $V_1 \ll V_0$ (i.e., quality requirements of the obligation are likely to not be met). In fact, it holds

$$\alpha_P(d_{01}) = a_P + \max\{0; 1 - a_P - d_{01}\} + \max\{0; 1 - 2a_P - d_{01}\} \tag{1}$$

where $a_P = 10\%$ and $d_{01} := V_1/V_0$. Namely, the base value of α_P is equal to 10%, but it is increased by 1% per each percentage point exceeding a 10% bidding discount and by 2% per each percentage point exceeding a 20% bidding discount.

Hence, in the case of $V_1 \simeq V_0$, $\alpha_B = 2\%$ and $\alpha_P = 10\%$, the exposure at risk guaranteed by a given performance bond is approximately five times the exposure at risk covered by the corresponding bid bond. However, depending on the choice of α_B and the value of d_{01}, the notional value N_P of the performance bond can easily reach an amount ranging between 10 and 20 times the corresponding bid bond exposure N_B. This fact incentivizes the insurer to assess the contractor as accurately as possible when underwriting the bid bond to avoid the choice between the bid bond payment and the issue of a performance bond that causes the exposure generated by the contractor to be too big for the contractor's worthiness.

The performance bond exposure decreases over time as the completion percentage of execution increases. The effective exposure can decrease up to 20% of the initial exposure in $t = t_1$. However, the insurer is usually unaware of the execution status: there is no obligation for the beneficiary or the principal to keep the insurer updated unless a claim is notified.

In case of a claim, the insurer indemnifies the beneficiary, and then the subrogation takes place: the insurer acquires the right to recover from the principal the amount indemnified to the beneficiary [19].

3. Sustainability of a Bid Bond

In this section, the risk-adjusted economic sustainability of a bid bond is investigated. Section 3.1 presents the bidder's perspective, while the risk appetite of the surety is discussed in Section 3.2. The procuring entity is protected by the mechanism described in Section 2, which rules the whole bidding process. In particular, Equation (1) fixes the notional value of the performance bond to cope with the performance risk of the winning bid. However, the tender process is sustainable for the procuring entity only if it leads to a sustainable outcome for both the contractor and the surety, such that the tender does not have to reopen. Hence, the sustainability conditions of all the subjects involved in the tender process are discussed in the following sections, explicitly or not.

3.1. The Bidder's Perspective

Let us assume that C is the contractor's cost to fulfill the principal obligation (e.g., the construction of public infrastructures). Inflation effects on the price of raw materials are negligible through the period $(t_0, t_1]$ when the tender takes place. Further, all the bidders are supposed to have access to the same liquid market to get the needed workforce, materials, and instruments. Thus, C is supposed to be equal to all the competitors and independent from time. We can safely assume that C also includes a minimum target profit required by each contractor's stakeholders. Its value remains (approximately) the same for all the bidders, even considering this additional contribution.

From a bidder's perspective, V_1 has both natural upper and lower bounds. The upper bound $\overline{V}_1 := V_0 > V_1$ holds true by construction, while the lower bound

$$\underline{V}_1 := C + \pi_P \leq V_1 \tag{2}$$

is the amount needed by the bidder to cover the expected costs, the target profit, and the risk-adjusted price $\pi_P(d_{01})$ of the performance bond. Filtered at t_1, the surety is assumed to price the performance bond according to a typical non-life actuarial pricing form [23,24] that accounts for expected losses, costs, and a prudential loading needed to compensate the surety's risk aversion, the latter also being the surety's profit. Hence, the price π_P obeys the equation

$$\pi_P = p_{12} L_P N_P + (r + s)\pi_P, \tag{3}$$

where p_{12} is the breach/default probability of the contractor over the execution period, $(t_1, t_2]$ and thus, $p_{12} L_P N_P$ is the expected loss contribution. Indeed, $L_P \in (0, 1]$ is the loss given default mitigation coefficient. It considers both the expected recovery of the surety from the contractor after the claim and the possible reduction of N_P at the claim time due to the partial fulfillment of the guaranteed obligation. The term $r\pi_P$ is the compensation for the risk aversion of the surety's stakeholders, where r scales as the cost-of-capital rate. The contract's price π_P is approximately proportional to the contract's contribution to the solvency capital requirement (SCR) needed by the insurer to guarantee its solvency in the Solvency II Standard Formula framework. Moreover, the surety's costs are assumed to be proportional to π_P and are taken into account by the cost ratio s term. The discount contribution of non-zero risk-free interest rates is neglected. As discussed in Section 2, it holds

$$N_P = \alpha_P(d_{01}) V_1, \tag{4}$$

where also $V_1 = d_{01} V_0$ is dependent on d_{01}. Hence, the risk-adjusted price of the performance bond is

$$\pi_P(d_{01}) = \frac{p_{12} L_P}{1 - r - s} \alpha_P(d_{01}) d_{01} V_0 \tag{5}$$

The price π_B of the bid bond follows the same structure and assumptions (see Remark 2 below).

Equations (2) and (5) imply that the public tender is *sustainable* for the winner of the tender, only if the inequality

$$\left[1 - \tfrac{p_{12} L_P}{1-r-s} \alpha_P(d_{01})\right] d_{01} \geq \frac{C}{V_0} \tag{6}$$

is verified. It is worth noticing that d_{01} is deterministic in the bidder's perpective, since it is a bidder's decision. On the other hand, p_{12} and r are unknown to the bidder, but a non-binding offer from the surety market is usually available on request, allowing the bidder to consider the provisional price

$$\hat{\pi}_P = \frac{\hat{p}_{12} L_P}{1-r-s} N_P(\hat{d}_{01}) \tag{7}$$

where $\hat{p}_{12}$ is the bidder's probability of default in $(t_1, t_2]$ estimated by the surety conditioned to the information available in t_0, and $\hat{d}_{01}$ represents the expected value of d_{01} under the same filtration $\mathcal{F}_{t_0}$. Hence, the boundaries $\overline{V}_1$ and $\underline{V}_1$ imply the determination of a compact interval where the contractor's choice of d_{01} in t_0 is rational.

$$\frac{C + \hat{\pi}_P}{V_0} \leq d_{01} \leq 1 \tag{8}$$

Condition (8) confirms two intuitive facts. First, the least risky bidder can afford to offer the greatest decrease of the starting price V_0, implying that the better the creditworthiness is, the higher the probability of winning the tender is. Second, the lower the ratio C/V_0 is, the smaller the minimum sustainable d_{01} value is.

Remark 1. *LHS of inequality* (8) *is a special case of inequality* (6)*, conditioned to the information available in t_0. Even if each bidder behaves rationally, placing a bid $d_{01}(t_0)$ in the interval defined in condition* (8)*, in t_1 it is still possible that the winner of the tender is awarded with a non-sustainable contract, because satisfying* (8) *in t_0 does not imply that* (6) *will be fulfilled in t_1. This uncertainty motivates the existence of prudential bids that are greater than the minimum rational level $\frac{C+\hat{\pi}_P}{V_0}$.*

Remark 2. *The cost π_B of the bid bond is negligible in the framework introduced above. In fact, it holds $\pi_B \ll \pi_P$, because $\alpha_B \ll \alpha_P$ (see Section 2) and $t_1 - t_0 \ll t_2 - t_1$. Further, the bid bond generates a claim only if the insured bidder defaults and is the winner of the tender. Hence, assuming to know the number N of participants involved in the tender process and considering approximately equal probabilities of being awarded among participants, the price of the bid bond can be written as*

$$\pi_B = \frac{1}{N} \frac{p_{01} L_B}{1-r-s} \alpha_B V_0. \tag{9}$$

under the same assumptions considered for π_P in Equation (5)*. As discussed above, the probabilities of winning are not uniform among the bidders, but such precise information is not available to a surety that guarantees just one of them in most cases. In general, considering the respective durations of bid bonds and performance bonds, and the $1/N$ factor as well, it holds*

$$\frac{1}{N} p_{01} \ll p_{12}, \tag{10}$$

that strengthens the validity of $\pi_B \ll \pi_P$.

The bid bond prices should be regarded more as "generic" expenses of the contractor than costs related to specific tenders, given that each contractor has to allocate a share of economic resources to participate in tenders, to win a part of them at most.

3.2. The Surety's Perspective

The subject who acts as the surety may be either a bank or an insurance company operating in the suretyship insurance line of business. We consider the latter case in the following, assuming that the Solvency II framework regulates the surety. This assumption copes with the investigated problem (i.e., the sustainability of bid bonds in Italy—a country where Solvency II is applied to the insurance market).

According to the Solvency II Directive [20] (Article 44), each insurer must define a set of rules, known as *Risk Appetite Framework* (also RAF), which aims to limit the capital absorption level below a given fraction of the own funds. This concept is then implemented in the Italian insurance law as well [25]. Since the RAF should discipline the business strategy and the management actions, the problem of the efficient capital allocation among the insurer's lines of business has been widely investigated in the actuarial literature (see, e.g., [26,27] and references therein).

However, this work focuses on the sustainability of a specific suretyship contract. Hence, our interest in a surety's RAF is limited to the subset of rules that may limit the surety's risk appetite against *Premium Risk* and the related *Catastrophe Man-Made Risk* in the Solvency II Standard Formula framework. On the other hand, the maximum acceptable amount of capital absorbed by the suretyship line of business is assumed to be fixed. Let us consider a (sub)portfolio composed of suretyship policies only. According to the Solvency II Standard Formula [21,22], such a portfolio exposes the insurer to three risk components of the *Underwriting Risk*:

i. The *Premium Risk*, whose Solvency Capital Requirement (SCR) is measured as

$$\begin{aligned} \mathrm{SCR}_{\mathrm{Pr}} &:= 3\sigma_{\mathrm{Pr}} V_{\mathrm{Pr}}, \\ V_{\mathrm{Pr}} &:= \max\{P_{\mathrm{Next}}; P_{\mathrm{Last}}\} + FP_{\mathrm{Existing}} + FP_{\mathrm{Future}}; \end{aligned} \tag{11}$$

where P_{Last} and P_{Next} are the premiums earned in the last 12 months and the premiums to be earned in the next 12 months, respectively; FP_{Existing} and FP_{Future} are the expected present value of the premiums to be earned after the following 12 months for existing contracts and for contracts whose initial recognition date falls in the following 12 months (for future contracts, premiums earned during the first 12 months after the initial recognition date are excluded from FP_{Future} contribution to volume measure), respectively; and $\sigma_{\mathrm{Pr}} = 19\%$ is the coefficient of variation associated to this sub-module of risk by the European regulator. The geographical diversification factor is not considered in Equation (11), since we are considering risks arising from Italian contractors only. The effect of reinsurance is ignored as well for this risk component and the next two listed below.

ii. The *Catastrophe Recession Risk*, whose Solvency Capital Requirement (SCR) is measured as

$$\mathrm{SCR}_{\mathrm{Rec}} := P_{\mathrm{Next}} \tag{12}$$

iii. The *Catastrophe Default Risk*, whose Solvency Capital Requirement (SCR) is measured as

$$\mathrm{SCR}_{\mathrm{Def}} := \mathrm{lgd}(LE_1 + LE_2) \tag{13}$$

where LE_i ($i = 1, 2$) are the first and the second largest exposures in the considered portfolio and lgd $= 10\%$ is a loss given default coefficient fixed by the European regulator.

The Standard Formula aggregation rule for the risk components listed above is

$$\begin{aligned} \mathrm{SCR}_{\mathrm{Udw}} &= \left(\mathrm{SCR}_{\mathrm{Pr}}^2 + 2\rho \mathrm{SCR}_{\mathrm{Pr}} \mathrm{SCR}_{\mathrm{Cat}} + \mathrm{SCR}_{\mathrm{Cat}}^2\right)^{\frac{1}{2}}, \\ \mathrm{SCR}_{\mathrm{Cat}} &= \left(\mathrm{SCR}_{\mathrm{Def}}^2 + \mathrm{SCR}_{\mathrm{Rec}}^2\right)^{\frac{1}{2}}, \end{aligned} \tag{14}$$

where $\rho = 25\%$ and $\mathrm{SCR}_{\mathrm{Udw}}$ is the Underwriting Risk measure under the assumption that all the risk components different from $i. - iii.$ are null, as futher specified in the following remark.

Remark 3. *Equation* (14) *measures only a part of the SCR_{Udw} that each suretyship insurance company has necessarily to cover. In particular, the Reserve Risk sub-module has been ignored, since this study is focused on the growth of Premium Risk due to newly underwritten contracts, which is directly related to the sustainability of the new policies.*

This simplification can be interpreted either as the assumption of instantaneous indemnifications (i.e., the surety opens and immediately closes the reserve provision associated with each claim, keeping the Reserve Risk negligible) or as the assumption that the surety's RAF disciplines *the Reserve Risk capital requirement separately from the Premium and Catastrophe risks. Indeed, the latter assumption is more likely than the first one.*

Lapse Risk is ignored since it is not considered relevant to this line of business.

Loosely speaking, in this context, the risk measure $\mathrm{SCR}_{\mathrm{Udw}}$ scales approximately with the size of the future earned premiums that, according to Equations (5) and (9), are proportional to the notional exposures ($N_B = \alpha_B V_0$ or $N_P = \alpha_P V_1$ in case of bid bonds or performance bonds, respectively) of each bond underwritten and to the corresponding claim probabilities ($\frac{1}{N} p_{01}$ or p_{02} respectively). Further, both N_B and N_P are proportional to the initial value V_0 of the contract and the performance bond exposure N_P has also a non-linear positive dependence on d_{01}, as shown in Equations (1) and (4).

For the sake of simplicity, let us consider a stable or expanding business, so that

$$\max\{P_{\mathrm{Next}}; P_{\mathrm{Last}}\} = P_{\mathrm{Next}}. \tag{15}$$

The simplification introduced in Equation (15) implies that Equation (14) can be rewritten as follows

$$\mathrm{SCR}_{\mathrm{Udw}} = \sqrt{9\sigma_{\mathrm{Pr}}^2 (P+F)^2 + 6\rho\sigma_{\mathrm{Pr}}(P+F)\sqrt{D^2+P^2} + D^2 + P^2} \tag{16}$$

where we use the compact notation $D := \mathrm{SCR}_{\mathrm{Def}}$ and $P := P_{\mathrm{Next}}$. Equation (16) allows to estimate the marginal contribution $\delta\mathrm{SCR}_{\mathrm{Udw}}$ to the capital requirement originated by a newly underwritten policy

$$\begin{aligned} \delta\mathrm{SCR}_{\mathrm{Udw}}(P,F,D) &= \mathrm{SCR}_{\mathrm{Udw}}(P+\delta P, F+\delta F, D) - \mathrm{SCR}_{\mathrm{Udw}}(P,F,D) \\ &= C_P\,\delta P + C_F\,\delta F + \ldots, \end{aligned} \tag{17}$$

where

$$C_P := \frac{1}{\mathrm{SCR}_{\mathrm{Udw}}(P,F,D)}\left[9\sigma_{\mathrm{Pr}}^2(P+F) + 3\rho\sigma_{\mathrm{Pr}}\frac{D^2+P^2+P(P+F)}{\sqrt{D^2+P^2}} + P\right], \tag{18}$$

$$C_F := \frac{1}{\mathrm{SCR}_{\mathrm{Udw}}(P,F,D)}\left[9\sigma_{\mathrm{Pr}}^2(P+F) + 3\rho\sigma_{\mathrm{Pr}}\sqrt{D^2+P^2}\right], \tag{19}$$

δP is the new policy's contribution to P_{Next}, and δF is the new policy's contribution to $FP_{\mathrm{Existing}} + FP_{\mathrm{Future}}$. D is assumed to be constant, which is generally true, unless the new policy's exposure exceeds LE_2 in Equation (13).

The premium accrual is linear in time, although the risk generated by the policy decreases as a non-linear function of the time-to-maturity. Hence, it holds that

$$\begin{aligned} \delta P &= \pi\frac{\min\{T-t,1\}}{T-t_0}, \\ \delta F &= \pi\frac{\max\{T-t-1,0\}}{T-t_0}, \end{aligned} \tag{20}$$

where π is the bond premium, t is the observation date, and t_0 and T are the recognition date and the maturity date of the bond, respectively. Abrupt variations of $\mathrm{SCR}_{\mathrm{Udw}}$ may occur in case an insured bidder wins a tender and the surety issues the performance bond as needed. In this case, the bid bond premium π_B is replaced by the corresponding performance bond premium $\pi_P \gg \pi_B$ (see Remark 2). The surety's RAF should aim to prevent the exposure from "jumping", associated with the conversions of bid bonds into performance bonds, which may lead to a breach of the established $\mathrm{SCR}_{\mathrm{Udw}}$ threshold level.

Many policy underwriters are simultaneously and independently operating on behalf of the surety. Hence, the contribution of each issued bond to $\mathrm{SCR}_{\mathrm{Udw}}$ cannot be taken into account instantaneously. In a realistic situation, $\mathrm{SCR}_{\mathrm{Udw}}$ is likely to be updated quarterly or twice a year, while new policies are issued daily or weekly. Hence, the surety may choose to maintain $\mathrm{SCR}_{\mathrm{Udw}}$ at a safe distance from a threshold $\overline{\mathrm{SCR}}_{\mathrm{Udw}}$ by defining a maximum acceptable $\delta\mathrm{SCR}_{\mathrm{Udw}}$ caused by each newly underwritten policy. Equations (5), (9), (17), and (20) imply that the maximum acceptable exposure $\overline{E}_p$ of a new policy scales as p^{-1}, where p is the claim probability of the policy. Namely, it holds

$$\overline{E}_p = \frac{\overline{\delta\mathrm{SCR}}_{\mathrm{Udw}}}{C_P \frac{\min\{T-t,1\}}{T-t_0} + C_F \frac{\max\{T-t-1,0\}}{T-t_0}} \frac{1-r-s}{L} p^{-1}, \tag{21}$$

where $\overline{\delta\mathrm{SCR}}_{\mathrm{Udw}}$ is the maximum acceptable variation of $\mathrm{SCR}_{\mathrm{Udw}}$ due to the new risk. C_P and F_P depend on the last updated values of SCR_{Udw}, P, F, and D. Terms beyond the first order in Equation (17) are assumed to be negligible.

Two concerns should be addressed before using Equation (21) to define a (simplified) surety RAF. First, if the contract is a bid bond, the case the contractor wins the tender and, thus, a performance bond is needed and must be considered. This issue is addressed later in Definition 2. Further, sureties want to limit their concentration of exposure against each contractor. Thus, a penalty term due to existing exposures that the same contractor generates should be considered.

To address the latter issue, the threshold $\overline{\delta\mathrm{SCR}}_{\mathrm{Udw}}$ is lowered by the first-order contribution to SCR_{Udw} of the policies already underwritten by the same contractor. Applying Equation (17) once again, we have the new threshold

$$\overline{\delta\mathrm{SCR}}_i := \max\left\{0, \overline{\delta\mathrm{SCR}}_{\mathrm{Udw}} - C_P \sum_{j\in\{i\}_t} \delta P_{ij} - C_F \sum_{j\in\{i\}_t} \delta F_{ij}\right\} \tag{22}$$

where $\{i\}_t$ is the sub-portfolio of policies existing in t and underwritten before t by the i-th contractor, and

$$\begin{aligned} \delta P_{ij} &= \pi_{ij} \frac{\min\{T^{(ij)}-t,1\}}{T^{(ij)}-t_0^{(ij)}}, \\ \delta F_{ij} &= \pi_{ij} \frac{\max\{T^{(ij)}-t-1,0\}}{T^{(ij)}-t_0^{(ij)}} \end{aligned} \tag{23}$$

are the contributions to P and F of the j-th policy in $\{i\}$, given the same notation used in Equation (20). It is worth noticing that the new threshold can be equal to zero, in case the concentration level on the i-th contractor has already exceeded the surety's risk appetite.

To handle the first concern on Equation (21), the surety's RAF can be defined as follows.

Definition 1 (Backward-looking Surety's RAF)**.** *The RAF is specified by the function* $\Psi : (0,1) \to \mathbb{R}_+$*, defined as follows.* $\Psi(p_{iJ})$ *is the maximum increment of exposure* $\overline{\delta E_i}$ *that the surety is allowed to guarantee against the i-th risky contractor (*$|\{i\}_t| = J-1$*), by issuing the new iJ-th bond whose claim probability is equal to* p_{iJ}*.* $\Psi : p_{iJ} \mapsto \overline{\delta E_i}$ *has the form*

$$\Psi(p_{iJ}) = \frac{\overline{\delta SCR_i}}{C_P \frac{\min\{T^{(iJ)}-t,1\}}{T^{(iJ)}-t_0^{(iJ)}} + C_F \frac{\max\{T^{(iJ)}-t-1,0\}}{T^{(iJ)}-t_0^{(iJ)}}} \frac{1-r-s}{L} p_{iJ}^{-1}. \tag{24}$$

Thus, the surety refuses to underwrite each J-th contract, such that $\delta E_{iJ} > \Psi(p_{iJ})$*.*

The RAF in Definition 1 is *backward-looking* in the sense that the acceptance or rejection of a given contract depends only on the contribution of the contract to the last SCR measured. As anticipated, Definition 1 does not offer solution to the first issued raised above (i.e., bid bonds which cope with Definition 1 may lead to performance bonds that exceed the frontier defined in Equation (24) in the future). Definition 2 also handles this issue.

Definition 2 (Forward-looking Surety's RAF)**.** *The RAF is specified by the couple* $\{\Psi(\cdot); p_\Psi\}$*. The function* $\Psi : (0,1) \to \mathbb{R}_+$ *defines the maximum increment of exposure* $\overline{\delta E_i}$ *that the surety is allowed to guarantee against the i-th risky contractor* ($|\{i\}_t| = J-1$) *by issuing the new* iJ*-th bond whose claim probability is equal to* p_{iJ}*.* $\Psi : p_{iJ} \mapsto \overline{\delta E_i}$ *has the form stated in Equation* (24)*. The tolerance* $p_\Psi \in (0,1)$ *is the maximum admissible probability that a risk underwritten in t implies a breach of the boundary* $\{p_{i'j'}, \overline{\delta E_{i'}}\}$ *at some* $t' > t$ *for all* i', j'*. Namely, the* iJ*-th bond can be underwritten in t only if*

$$\mathbb{P}\left[\Psi\left(p_{i'j'}(t')\right) < E_{i'}(t')\Big|\mathcal{F}_t\right] < p_\Psi \quad \forall t' > t,\, i', j' \in \mathbb{N}, \tag{25}$$

where $\mathbb{P}$ *is real-world probability measure available to the surety. A bond that satisfies both conditions* (24) *and* (25) *by t is* sustainable *in the surety's perspective.*

The Solvency II Standard Formula is based on some simplifying assumptions that also affect Equation (24).

First, premiums as a volume measure establish a link between the riskiness of each risk source (i.e., the contractor in this case) and the capital requirement. However, in case the premium rate is fixed at the issuing date t_0, it is related to the contractor's standing at t_0, but it may not be representative of the contractor's riskiness when SCR_{Udw} is evaluated.

Further, the fraction of premium to be earned by the surety decreases linearly over time. Hence, the negative dependence between risk and residual time-to-maturity is taken into account. However, the non-linear decreasing of risk by time—as shown, e.g., in Equation (7)—is replaced by a linear dependency.

Despite these limitations, the Standard Formula represents a breakeven point between simplicity and effectiveness. Being an established standard in the European insurance industry, it is worth considering it when defining the RAF used to investigate the sustainability of a given bid bond. Internal model approaches are possible as well and are not affected by the limitations mentioned above. However, in this work, we are interested in investigating possible paradoxes arising in a standard context. Hence, we chose to use the Standard Formula exclusively.

Definition 1 introduces a maximum exposure-at-risk $\overline{E}_i$ per contractor, implicitly. The value of $\overline{E}_i$ depends on the standing of the i-th contractor and the remaining time-to-maturity of each underwritten contract, in agreement with intuition. On the other hand, Definition 2 also forbids less trivial cases.

A bid bond that satisfies condition (24) in t_0 may still not comply with condition (25), in case it holds

$$\mathbb{P}\left[\Psi^{-1}(E_{iJ+1}(t_1)) < p_{iJ+1}(t_1)\Big|\mathcal{F}_{t_0^{(iJ)}}\right] > p_\Psi, \tag{26}$$

where the bid bond is the J-th policy underwritten with the i-th contractor and the subsequent performance bond (in case the contractor wins the tender) is the $J+1$ policy.

It is worth noticing that Equation (25) implies restrictions stronger than the one stated in Equation (26). Let us consider the iJ-th bid bond mentioned above, assuming that it satisfies condition (24) and does not have the problem in Equation (26). Even in such a case, the bond could still not satisfy condition (26) due to portfolio issues. In fact, when the number of simultaneously active bid bonds is large enough, the probability that one of them results in a future performance bond not compliant with condition (24) exceeds p_Ψ, even if the last underwritten bid bond complies with the RAF when considered stand-alone.

The "global" sustainability of a bid bond (i.e., in the context of the surety's portfolio of underwritten bonds) is addressed in Section 4 numerically.

4. Measuring and Managing the Unsustainability Scenarios in Public Tenders

This section addresses the sustainability issues introduced in Section 3 by implementing the surety's RAF proposed in Definitions 1 and 2. A model is introduced in Section 4.1 to simulate the tenders. Each simulation considers three alternate versions of the surety: without an RAF, adopting a backward-looking RAF as per Definition 1, and adopting a forward-looking RAF as per Definition 2. The results obtained by the Monte Carlo simulations are presented in Sections 4.2–4.4, respectively.

4.1. Simulation of Tenders from a Surety's Perspective

In the following paragraphs, the model employed to simulate the tenders of public works is described. The model aims to highlight that unsustainable requests for a performance bond are possible when considering realistic dynamics of default (or breach) probabilities associated with each bidder. Further, the model is employed to investigate the effectiveness of the RAF strategies implemented by the surety.

A realistic model of the Italian public works market should include some aspects not considered in this section, such as the actual number of tenders per year where the considered surety guarantees at least a bidder and the distribution of the public works costs C. However, the data needed to calibrate such a model are non publicly available, and the model itself would not fit better for the purpose of this study than the toy model introduced here.

Let us consider a where the elementary time step δt is a quarter long. We consider a surety with access to 10^3 public tenders per quarter, issuing a bid bond to at least a participant per tender since there is no need to simulate tenders where the considered surety has no business. We assume that the initial price of each tender is a uniform r.v.

$$V_0 \sim \text{Unif}_{[C,\,150\%C]}. \tag{27}$$

Both the boundaries $\min V_0$ and $\max V_0$ are admitted to represent a possible misjudgment of the procuring entity. In fact, $V_0 = C$ leaves no room to lower the initial price or to aim for an extra profit, implying that no bidder is joining the tender. Further, in case $V_0 = 150\%\,C$, Equation (1) implies that E_P/E_B can reach a value of 45 and above, increasing the probability that the required performance bond violates Equation (24) and, thus, that sureties reject the (unsustainable) winner's request for a performance bond.

Thus, it is natural to assume a positive dependency between V_0/C and the number of potential bidders $\tilde{N}$ interested in joining the tender. We chose the form $\tilde{N} = \left\lfloor 100\left(e^{V_0/C-1} - 1\right)\right\rceil$, which implies a realistic range $\tilde{N} \in \{0 = N_{V_0=C}, \ldots, N_{V_0=1.5\cdot C} = 65\}$. However, the $N \leq \tilde{N}$ constructors who actually join the tender are the ones able to make a bid according to the condition (8), depending on the values of $\hat{\pi}_P$ and V_0.

The surety can issue more than a bid bond per tender (up to N), increasing the probability that one among its insured bidders wins the contract and is required to underwrite the corresponding performance bond. However, considering the competition in the surety market, we assume that the number of bidders n joining the same tender and insured by the same surety is distributed as a shifted Poisson r.v. Namely, $\tilde{n} - 1 \sim \text{Pois}(\lambda_B)$, where $\lambda_B = 0.1$ and $n = \min\{\tilde{n}; N\}$.

The parameter α_B is assumed to be distributed as a categorical r.v. with probability mass function

$$f(\alpha_B) = \begin{cases} 0.2, & \alpha_B \in \{1\%; 3\%; 4\%\}; \\ 0.4, & \alpha_B = 2\%; \\ 0.0, & \text{otherwise.} \end{cases} \tag{28}$$

where the mode is fixed at 2%, as anticipated in Section 2.

To estimate claim probabilities and their dynamics, we consider historical time series of performing ("PL") and non-performing loans ("NPL") [28], publicly available from the Bank of Italy [29]. This choice is justified by the assumption that the claim probability of a contractor is completely correlated to its creditworthiness. This is true in the extreme case of bankruptcy, which implies the contractor's inability to be operating. In general, it is a fair approximation, although other elements of technical nature (e.g., unforeseen geological features of the building location) may contribute to the performance risk in specific cases.

Time series PL_t and NPL_t are quarterly available by ATECO 2007 economic sector (i.e., our data are restricted to the "constructors" sector), size of loan s (three clusters) and geographical location g of the Italian debtor (five clusters). Hence, dynamics of claim probability can be specified by considering 15 bivariate time series $\{\mathrm{PL}_t; \mathrm{NPL}_t\}_{sg}$, where PL_t is the number of performing loans at the first day of the t-th quarter, while NPL_t is the number of loans that become past due during the t-th quarter.

Since we need to represent a significant number of contractors by introducing a parsimonious number of parameters, we choose to apply the CreditRisk$^+$ model [30,31] to describe the dependence structure among the claims and the marginal volatility of each cluster probability of default. The CreditRisk$^+$ model defines the dependence among defaults (or other absorbing events, such as breaches of contracts) through an array of latent market factors $\mathbf{\Gamma} \in \mathbb{R}_+^K$, where $\Gamma_k \sim \mathrm{Gamma}(\sigma_k^{-2}, \sigma_k^2)$, $(k = 1, \dots, K)$. It holds that $\mathbf{E}[\Gamma_k] = 1$ and $\mathrm{Var}[\Gamma_k] = \sigma_k^2$ by construction. The market factors alter the parameter's value of the r.v. $Y_i(t, t')$, which represents the occurence of a claim generated by the i-th contractors in the time interval $(t, t']$. In its original formulation, the model [30] is defined in a single-time-scale framework and $Y_i \dot{\sim} \mathrm{Pois}(p_i)$, where

$$p_i(\mathbf{\Gamma}) \coloneqq q_i \cdot \left(\omega_{i0} + \sum_{k=1}^{K} \omega_{ik} \Gamma_k\right) \tag{29}$$

and the factor loadings ω_{ik} are supposed to be all non-negative and to sum up to unity:

$$\begin{aligned} &\omega_{ik} \geq 0, \qquad i = 1, \dots, N, \quad k = 0, \dots, K, \\ &\sum_{k=0}^{K} \omega_{ik} = 1, \quad i = 1, \dots, N. \end{aligned} \tag{30}$$

We consider the model's generalization recently proposed in [32], which has also been applied to credit and suretyship insurance [33]. The advantage of this choice is the possibility to calibrate the model by using the quarterly time series available and using it to estimate both the bid bond claim probabilities $p_{01}(t, h(i))$ and the corresponding performance bond claim probabilities $p_{12}(t, h(i))$, where $h = 1, \dots, 15$ labels the cluster of the i-th bidder.

We assume that each bid bond has a 3-month coverage period on the interval $(t = t_0, t_1 = t + \delta t]$, while each performance bond has a 5-year coverage period on the interval $(t_1, t_1 + 20\delta t]$. This simplification is a part of the toy framework that we are defining since the duration of public works depends on each project's features and size. However, it is numerically sound since a tender process takes a few months to close, while a public works project typically lasts a few years. Hence, a two-time-scales parameterization is needed to price both the bid and the performance bond.

According to [32], the claim event in CreditRisk$^+$ framework can be modeled as $Y_i(t, t') \sim \mathrm{Bernoulli}(p_i(t, t'))$, where the parameter p_i has an exponential dependency on the latent factors. Namely, under our set of assumptions, it holds that

$$p_{01}(t_0, h) = 1 - \exp\left[-q_h \left(\omega_{h0} + \sum_{k=1}^{K} \omega_{hk} \Gamma_k(t_0)\right)\right], \tag{31}$$

$$p_{12}(t_1, h) = 1 - \exp\left[-20 q_h \left(\omega_{h0} + \frac{1}{20} \sum_{\tau=0}^{19} \sum_{k=1}^{K} \omega_{hk} \Gamma_k(t_1 + \tau\, \delta t)\right)\right], \tag{32}$$

where $\tau \in \mathbb{N}$ is the index used to label each quarter. Further, assuming that the surety has developed an internal rating model such that reliable estimates of

$$\{\boldsymbol{\Gamma}(t) : t = t_0, \ldots, t_0 + 19\delta t\} | \mathcal{F}_{t_0},$$

we can represent the possible fluctuations in performance bond pricing by considering $\hat{p}_{12} = p_{12}(t_0, h(i)) = p_{12}(t_1 - \delta t, h(i))$, while the (correct) estimate p_{12}, which allows the computation of π_P, shall not be available until $t = t_1$.

Let us consider the time series $\{\mathrm{PL}_t; \mathrm{NPL}_t\}_{sg}$ available in [29], from the first quarter of 2016 to the first quarter of 2021, to calibrate the CreditRisk$^+$ model. The generalized covariance estimator defined in [32], the decomposition technique introduced in [34], and the standard regularization technique described in [35] return the result in Table 2.

Table 2. The complete set of parameters $\hat{\Omega}, \hat{\sigma}^2_\Gamma$ necessary to specify the dependence structure in CreditRisk$^+$ model applied to the Italian "*Constructors*" economic sector.

s	g	h	q_h	$\mathbf{k} =$	**0**	**1**	**2**	**3**	**4**	**5**
1	North-West	1	0.0050		0.516	0.246	0.025	0	0.212	0
1	South	2	0.0095		0.473	0.220	0.019	0.216	0	0.072
1	Islands	3	0.0098		0.559	0.231	0	0.210	0	0
1	North-East	4	0.0040		0.658	0.254	0.088	0	0	0
1	Center	5	0.0079		0.026	0.234	0.070	0.432	0	0.238
2	North-West	6	0.0079		0.509	0.228	0	0.028	0.235	0
2	South	7	0.0137		0.651	0.292	0.057	0	0	0
2	Islands	8	0.0158	$\omega_{hk} =$	0.455	0.266	0	0.036	0.237	0.005
2	North-East	9	0.0073		0.628	0.281	0.086	0.006	0	0
2	Center	10	0.0126		0.590	0.281	0.057	0.016	0	0.056
3	North-West	11	0.0130		0.091	0.235	0	0.134	0.443	0.096
3	South	12	0.0147		0.467	0.236	0.048	0	0.127	0.122
3	Islands	13	0.0185		0.619	0.333	0	0	0	0.049
3	North-East	14	0.0134		0.611	0.265	0.050	0	0.068	0.005
3	Center	15	0.0178		0.325	0.251	0.099	0	0	0.326
				$\sigma^2_k =$		2.417	0.157	0.052	0.049	0.044

Remark 4. *Claims are not simulated in our setting, although the CreditRisk$^+$ framework is explicitly designed to do it. In fact, they are not relevant to the part of the surety's RAF addressed in this study. Occurred claims affect mainly the reserve provision and the Reserve Risk capital requirement. On the other hand, they may lead to a slight decrease of SCR_{Cat} or SCR_{Pr}, since claims generated by suretyship insurance products are* absorbing events. *Each policy may generate one claim at most during the coverage period, implying the zeroing of both the corresponding exposure and future premiums (if any).*

The framework is completed by associating each i-th bidder to its $h(i)$-th cluster. This is achieved by modeling the position of the public works underlying each tender as a categorical random variable. The probability associated with each g-th area ($g = 1, \ldots, 5$) is proportional to the number of performing borrowers belonging to that area observed in the construction sector by the first quarter of 2021. This is equivalent to assuming a correspondence between demand and offer in this economic sector (i.e., the presence of many constructors implies a relevant number of public tenders and vice versa). Assuming that all the bidders belong to the same area where the public works must be executed, their distribution among the three loan classes is modeled in a similar way, considering a categorical variable per geographical area g, where probabilities are proportional to the number of performing borrowers observed in the cluster sg ($s = 1, 2, 3$), conditioned to g.

It is worth recalling that the public works tenders in Italy can be classified as first-price, sealed descending bid auctions (for complete classification of auctions and a deep theoretical discussion, see, e.g., [36,37]). The bid domain is compact, and the winner is the author of the lowest bid in a set of non-identically distributed bids. Hence, we cannot use the Fisher–Tippett–Gnedenko theorem [38,39], which is commonly applied to model the distribution of the winning bid in the ascending bid auctions (see, e.g., the recent paper [40], where a Weibull distribution is considered). Thus, we need to determine the winning bid numerically, considering that each rational bidder chooses its d_{01} according to condition (8).

A non-uniform distribution is assumed over each i-th domain to take into account the appetite of each bidder for obtaining the contract. Namely, it holds that

$$\frac{d_{01}V_0 - C - \hat{\pi}_P(i)}{V_0 - C - \hat{\pi}_P(i)} \sim \text{Beta}(\alpha, \beta) \tag{33}$$

where multiple specifications have been tested for the parameters set (α, β), as represented in Figure 3: mode[pdf(d_{01})] tends to $(C + \hat{\pi}_P)/V_0$ at increasing bidder's appetite for winning the tender, while it tends to V_0 at increasing appetite for profit. However, no relevant effect of the (α, β) choice is observed on the results presented in Sections 4.2–4.5. Hence, only the choice $(\alpha, \beta) \equiv (1.80, 7.20)$ is considered in Figures 4–7.

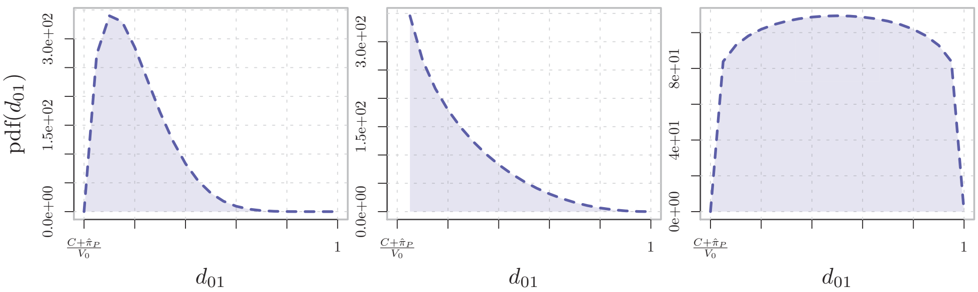

Figure 3. Different behaviors of the bidders modeled by alternative parameterizations of the Beta distribution in Equation (33). *Left panel*: $\alpha = 1.80$, $\beta = 7.20$. *Central panel*: $\alpha = 0.80$, $\beta = 3.20$. *Right panel*: $\alpha = 1.16$, $\beta = 1.16$.

The winner of each tender is the lowest simulated bid per tender/scenario. The bidders are indexed in simulations ($i = 1, \ldots, N$). In doing so, the case where the winning bidder is among the ones guaranteed by the considered surety is explicitly represented.

4.2. Dynamics of the Capital Requirement without Taking Management Actions

Let us consider a suretyship insurance company that operates as a surety in the framework introduced in Section 4.1. The surety is supposed to start operating in $t = 0$. It is worth recalling that the duration of a bid bond is established to be equal to three months, while each performance bond is assumed to expire after five years. Thus, as expected, the surety SCR_{Udw}—as defined in Equation (16)—reaches an equilibrium after five years, considering a stable flow (on average) of new contracts per year (Figure 4).

Without loss of generality, we choose $\overline{SCR}_{\text{Udw}} = 0.7C$, which is below the equilibrium level $\text{SCR}_{\text{Udw}}(t > 5) \approx 0.75C$ obtained numerically. The surety is supposed to increase its sales volume until its risk appetite level is reached. Then, a risk appetite framework is introduced to discipline the underwriting process, as discussed in Section 3.2. Hence, in a liquid market, a surety with a higher risk appetite or a larger amount of its own available funds than the one considered in our simulations would reach the same equilibrium state at a different SCR_{Udw} level.

4.3. Dynamics of the Capital Requirement Adopting a Backward-Looking RAF

In Section 4.2, the surety has reached an equilibrium state that is slightly above its risk appetite level. Thus, an RAF is needed to prevent the occurrence of breaches

$\mathrm{SCR_{Udw}}(t) > \overline{\mathrm{SCR}}_{\mathrm{Udw}}$. The same simulations presented in Figure 4 are re-performed, applying the management actions implied by Definition 1, *ceteris paribus*.

The level $\overline{\delta \mathrm{SCR}}_{\mathrm{Udw}}$ is needed to specify $\Psi(\cdot)$. It has to be as high as possible to refuse the minimum number of contracts per unit of time, conditioned to avoid breaches or, at least, make them improbable enough according to the surety's risk appetite. In the example, we chose

$$\overline{\delta \mathrm{SCR}}_{\mathrm{Udw}}(t) = \max\left\{0; \min\left\{\overline{\mathrm{SCR}}_{\mathrm{Udw}} - \mathrm{SCR}_{\mathrm{Udw}}(t - \delta t); \delta \mathrm{SCR}_{\mathrm{Udw}}\left[\Phi^{-1}_{\pi_P}(0.95)|\mathcal{F}_t\right]\right\}\right\}$$

where $\Phi_{\pi_P}(\cdot)$ is the cumulative distribution function of the performance bond prices π_P and $\delta \mathrm{SCR}_{\mathrm{Udw}}[\cdot|\mathcal{F}_t]$ is the marginal contribution of a given contract underwritten in t to $\mathrm{SCR_{Udw}}$. $\delta \mathrm{SCR}_{\mathrm{Udw}}$ is evaluated by applying the linear approximation stated in Equation (17).

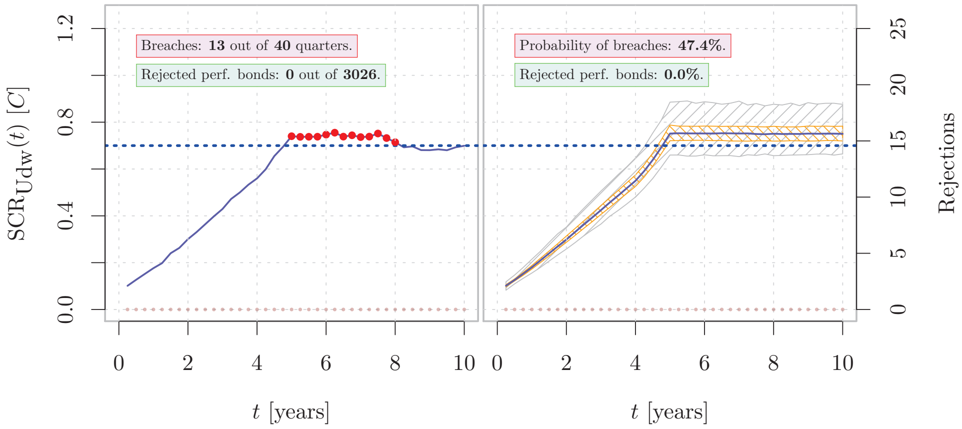

Figure 4. Dynamics of $\mathrm{SCR_{Udw}}(t)$. *Left panel*: example of Monte Carlo simulation (single scenario), where the horizontal dashed line represents the $\overline{\mathrm{SCR}}_{\mathrm{Udw}}$ threshold that is fixed in the RAF, while red dots correspond to the simulated breaches of $\mathrm{SCR_{Udw}}(t)$ above the $\overline{\mathrm{SCR}}_{\mathrm{Udw}}$ level. Red columns at 0 level (right y-axis scale) represent the count of rejected performance bonds per quarter (zero since no management action is taken). *Right panel*: simulated distribution of $\mathrm{SCR_{Udw}}(t)$ (10^3 Monte Carlo scenarios). The median (solid line), 0.25–0.75 quantiles (orange area), and 0.01–0.99 quantiles (grey area) are plotted. Red columns at 0 level (right y-axis scale) represent the average number of rejected performance bonds per quarter.

Namely, when the last measure $\mathrm{SCR_{Udw}}(t - \delta t)$ done until t is far enough from the threshold $\overline{\mathrm{SCR}}_{\mathrm{Udw}}$, we aim not to reject more than 5% of the performance bond requested by the insured bidders who win their respective tenders. In case the distance $\overline{\mathrm{SCR}}_{\mathrm{Udw}} - \mathrm{SCR_{Udw}}(t - \delta t)$ approaches zero or negative values, the RAF constraint becomes stronger up to blocking the acquisition of new contracts at all, until an acceptable $\mathrm{SCR_{Udw}}$ level is restored. Figure 5 shows the effectiveness of this approach. Breaches are observable in tail scenarios, almost only in the region ($t \simeq 5$) where the SCR regime is changing from *expansion* to *equilibrium*. The small number of breaches and the RAF's reaction implies a fraction of performance bonds rejected slightly above the 5% target level.

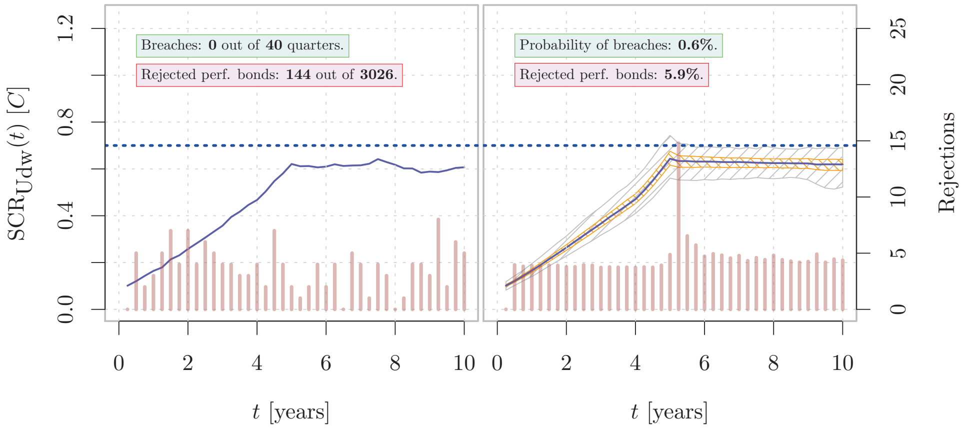

Figure 5. Dynamics of $\mathrm{SCR}_{\mathrm{Udw}}(t)$, given the notation introduced in Figure 4. RAF introduced in Definition 1 is applied to estabilish the management actions taken.

4.4. Dynamics of the Capital Requirement Adopting a Forward-Looking RAF

As discussed above, the surety should avoid the rejection of a performance bond as much as possible. This case opens the possibility that the tender must be reopened if no other surety is willing to issue the performance bond instead of the surety that has issued the bid bond to the winning contractor. Further, in a market where sureties are comparable, the rejection of a request by a company implies that the other companies in the same market are likely to do the same, leading to a claim generated by the bid bond.

Hence, it is worth addressing this issue by implementing the forward-looking RAF introduced in Definition 2. It is not necessary to establish p_Ψ explicitly. The probability that a performance bond generates a breach is an increasing function of V_0/C. Loosely speaking, a higher starting price implies that the winning bid—always close to C—corresponds to a greater discount $1 - d_{01}$. Thus, a higher α_P can be expected as well, increasing the probability of a breach $\overline{\delta \mathrm{SCR}_{\mathrm{Udw}}}$.

Even without knowing the analytical form of the dependencies described above, the qualitative picture is enough to implement the constraint (25) as

$$\Phi^{-1}_{V_0/C}(0.75) > \left(\frac{V_0}{C}\right)_k \tag{34}$$

where $\Phi_{V_0/C}(\cdot)$ is the cumulative distribution function of $\frac{V_0}{C}$, and $\left(\frac{V_0}{C}\right)_k$ is the ratio observed in the k-th tender. The percentile 0.75 has been chosen numerically with the aim of minimizing both the number of rejected performance bond requests and the frequency of SCR breaches. Bid bonds that do no cope with Equation (34) are rejected, preventing a possible unsustainable request for a performance bond (in case the bidder wins).

Given the additional constraint, we can weaken the other introduced in Section 4.3, passing from 0.95 to 0.99 (i.e., we aim to reject up to 1% of performance bonds, to limit both the claims arising from the corresponding bid bonds and the surety's reputational risk). Results are exposed in Figure 6, where the number of rejected performance bonds decreases.

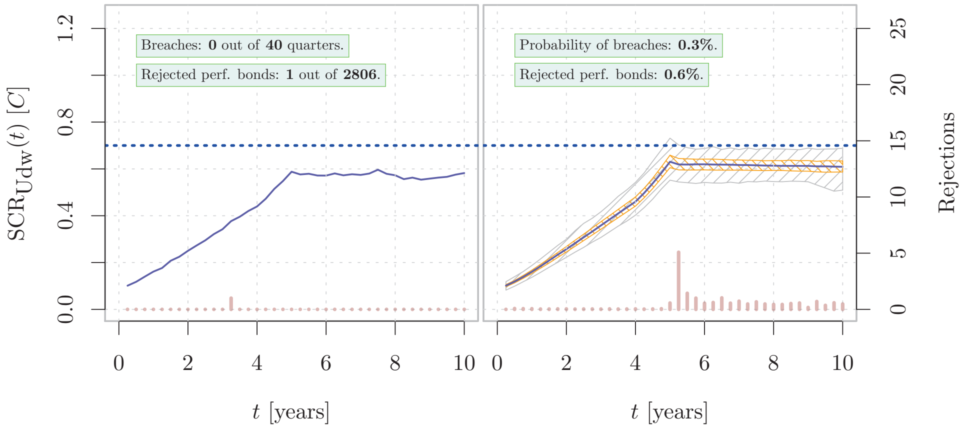

Figure 6. Dynamics of $\text{SCR}_{\text{Udw}}(t)$, given the notation introduced in Figure 4. RAF introduced in Definition 2 is applied to estabilish the management actions taken.

4.5. The Role of the Procuring Entity

As shown in Section 4.4, a surety that implements a forward-looking RAF can prevent the majority of unsustainable tender outcomes, avoiding rejecting performance bond requests from the winners and the possible subsequent need for reopening the tender.

However, unsustainability issues originate from a poor choice of starting price by the public procuring entity. The left panel of Figure 7 shows how a starting price near the breakeven level (i.e., $V_0 \simeq C$) disincentivizes constructors to join the bid due to the constraints introduced in condition (8). It is worth noticing that this result is independent of the assumption made about the dependency $\tilde{N}(V_0)$. A tender that does not attract participants is clearly unsustainable from an economic perspective. The resources invested in promoting it are wasted, and the public works cannot be executed. Further, the opposite case (i.e., $V_0 \gg C$) also implies the economic unsustainability of the tender, as the tender outcome implies $N_P/N_B \gg 1$ and thus an excessive risk for the surety (i.e., an unsustainable cost for the winning bidder or the inability to underwrite the mandatory performance bond). The right panel of Figure 7 shows the results of our simulations in this perspective: the fraction of requests for a performance bond rejected by a given surety increases from ≈0% to ≈50% as V_0 passes from ≈1.25C to ≈1.45C.

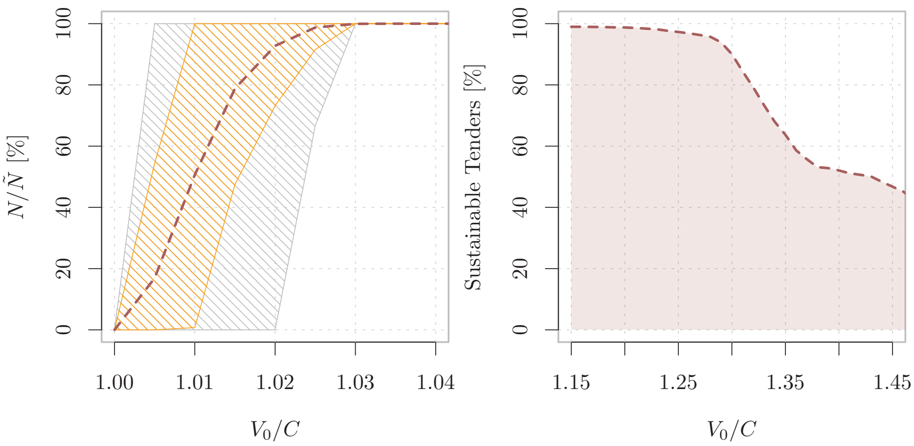

Figure 7. Sustainability as a function of the starting price V_0. *Left panel*: fraction of constructors who can afford to make a bid, as the starting price approaches the breakeven level C (*dashed line*: average; *orange area*: confidence interval within ± 1 standard deviation; *grey area*: confidence level within 0.01–0.99 quantiles). *Right panel*: fraction of concluded tenders whose requested performance bond is not rejected by the surety (average level) at increasing starting prices.

Remark 5. *Figure 7 shows that the tender is almost surely sustainable, depending on the procuring entity's proper choice of the V_0 value. This fact implies that the case $\pi_P \gg \hat{\pi}_P$, due to the worsening of a bidder's creditworthiness during the tender process, has a negligible impact. The numerical evidence copes with the following intuition: since the tender lasts a few months, a relevant change in a bidder's credit standing is unlikely during such a short period.*

5. Conclusions

This study highlights the existence of an economic unsustainability risk for the tenders that award public works contracts. The Italian bidding law and the Solvency II regulatory framework are explicitly considered to model the behavior of the three subjects involved in the tender process: the procuring entities, the bidders, and the sureties. Numerical simulations show that this risk can be mitigated and prevented by the proper choices of both the surety and the procuring entity.

In Section 3, sustainability conditions are stated for both the bidders and the sureties. In particular, sureties can protect their SCR target levels by applying an RAF, as requested by the Solvency II Directive. In Section 3.2, we proposed two simplified RAFs, both based on the linearization of the Solvency II Standard Formula. The first one (i.e., "backward-looking"—Definition 1) aims only to protect the surety's SCR level, regardless of the effects on the tender process. A boundary $(p; \overline{E}_p = Kp^{-1})$ is shown to separate sustainable new exposures $E \leq \overline{E}_p$ from the unsustainable ones ($E > \overline{E}_p$), depending on the default probability p associated to the considered contractor. A closed-form expression for K is provided from the Standard Formula prescription to evaluate the non-life underwriting risk module for the S2LoB 9-*Credit & Suretyship Insurance*. We show that the sureties can protect both themselves and the tender process by applying a "forward-looking" RAF, such as the one proposed in Definition 2. A numerical comparison among the two RAFs was presented in Sections 4.3 and 4.4, showing that the latter can actually protect all the three subjects involved in the tender process.

While the surety can mitigate the unsustainability risk, the public procuring entity can prevent it by establishing an acceptable starting price for the tender. Our simulations suggest that $V_0/C \in [1.05, 1.25]$ is the most sustainable choice in a realistic setting, where the Italian constructors' default probability is modeled by applying the CreditRisk$^+$ framework to recent historical data.

The proposed framework can be further investigated and improved in future studies. In particular, the Standard Formula approach can be replaced by a Partial Internal Model to define the two RAFs. The specific features of other regulatory frameworks diverse from the Italian bidding law may be investigated as well, provided that historical information to calibrate the model is publicly available for each considered country. Further, two limitations of this study are reported as follows, which can be addressed in further studies as well: First, both the procuring entity and the surety are assumed to perform error-free estimates. The procuring entity could actually choose V_0 poorly because of a bad strategy or the error affecting its C measure. The surety may perform a poor estimate of p_{12} as well, implying the missed identification of an unsustainable tender outcome due to a mistaken π_P evaluation. A second limitation of the study arises from the simplifications made in simulating the solvency balance sheet of the surety. Our model could consider a dynamical reserve risk and a non-zero market risk (generated by the surety's assets) in addition to premium and catastrophe risk components to provide a more realistic representation. Although future studies can adequately address these limitations, it is worth noticing that our simplified framework is consistent with the features that the considered system should have according to [16,17], as summarized in Section 1. These limitations do not diminish the practical conclusions of the study about the strategies that the surety and the procuring entity can implement to mitigate the investigated unsustainability risk, namely the adoption of the RAF introduced in Definition 2 and the choice of $V_0/C \in [1.05, 1.25]$, respectively.

Author Contributions: Conceptualization, J.G.; methodology, J.G.; software, J.G.; validation, L.P.; formal analysis, J.G. and L.P.; investigation, J.G.; resources, J.G. and L.P.; data curation, J.G. and L.P.; writing—original draft preparation, J.G.; writing—review and editing, L.P.; visualization, J.G. and L.P.; supervision, L.P.; project administration, J.G. and L.P.; funding acquisition, not applicable. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data that support the findings of this study are openly available in "Banca d'Italia—Base Dati Statistica" [29].

Conflicts of Interest: The authors declare no conflict of interest. The views and opinions expressed in this article are those of the author and do not necessarily reflect the official policy or position of SACE S.p.A.

References

1. Luque González, A.; Coronado Martín, J.Á.; Vaca-Tapia, A.C.; Rivas, F. How Sustainability Is Defined: An Analysis of 100 Theoretical Approximations. *Mathematics* **2021**, *9*, 1308. [CrossRef]
2. Directive 2014/24/EU of the European Parliament and of the Council of 26 February 2014 on Public Procurement and Repealing Directive 2004/18/EC. Available online: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex:32014L0024 (accessed on 1 September 2021)
3. "*Miller Act*", 24 August 1935. 49 stat. 793–794, ch. 642, Sec. 1–3. Codified as Amended in the *Code of Laws of the United States of America*, Title 40 *Public Buildings, Property, and Works*. Available online: https://www.law.cornell.edu/uscode/text/40 (accessed on 1 September 2021)
4. People's Republic of China-Ministry of Commerce *Bids Law of the People's Republic of China*, order No.21 of the President of People's Republic of China on 30 August 1999. Available online: http://english.mofcom.gov.cn/article/policyrelease/Businessregulations/201303/20130300046070.shtml (accessed on 1 September 2021)
5. Federal Law No. 223-FZ. *On Procurement of Goods, Works and Services by Certain Types of Legal Entities*; Signed into Law by the Russian President on 18 July 2011. Available online: http://government.ru/docs/all/100084/ (accessed on 1 September 2021)
6. Swiss RE. Trade credit Insurance & surety: Taking stock after the financial crisis. *Econ. Res. Consult.* **2014**. Available online: https://www.swissre.com/institute/research/topics-and-risk-dialogues/economy-and-insurance-outlook/expertise-publication-trade-credit-insurance-surety.html (accessed on 1 September 2021)
7. Glen Boswall, R. *Construction Bonds Guide*; Clark Wilson LLP: Vancouver, BC, Canada, 2010. Available online: https://www.cwilson.com/app/uploads/2010/09/construction-bonding-guide.pdf (accessed on 1 September 2021).
8. Russell, J.S. *Surety Bonds for Construction Contracts*, 1st ed.; American Society of Civil Engineers-ASCE Press: Reston, VA, USA, 2000.
9. Yesilyaprak, M.; Erkök, B. A new instrument in Turkish financial market: Surety bonds. *Finansal Araştırmalar ve Çalışmalar Dergisi* **2021**, *13*, 904–917. [CrossRef]
10. Choudhry, R.M.; Iqbal, K. Identification of Risk Management System in Construction Industry in Pakistan. *J. Manag. Eng.* **2013**, *29*, 42–49. [CrossRef]
11. Liu, J.; Li, B.; Lin, B.; Nguyen, V. Key issues and challenges of risk management and insurance in China's construction industry: An empirical study. *Ind. Manag. Data Syst.* **2007**, *107*, 382–396. [CrossRef]
12. Renault, B.Y.; Agumba, J.N. Risk management in the construction industry: A new literature review. *MATEC Web Conf.* **2016**, *66*, 00008. [CrossRef]
13. Bezer, D. The Inadequacy Of Surety Bid Bonds In Public Construction Contracting. *Public Contract Law J.* **2010**, *40*, 87–146. Available online: http://www.jstor.org/stable/25755803 (accessed on 1 September 2021).
14. Hassan, A.A.; Adnan, H. The problems and abuse of performance bond in the construction industry. *IOP Conf. Ser. Earth Environ. Sci.* **2018**, *117*, 012044. [CrossRef]
15. Helw, A.A.; Ezeldin, A.S. Toward Better Applicability of Public Procurement Law: Delay Claims by the Contractor and Limit of Compensation under the Performance Guarantee. *J. Leg. Aff. Disput. Resolut. Eng. Constr.* **2021**, *13*, 04521023. [CrossRef]
16. Calveras, A.; Ganuza, J.J.; Hauk, E. Wild bids. Gambling for resurrection in procurement contracts. *J. Regul. Econ.* **2004**, *26*, 41–68. [CrossRef]
17. Wambach, A.; Engel, A.R. Surety Bonds with Fair and Unfair Pricing. *Geneva Risk Insur. Rev.* **2011**, *36*, 36–50. [CrossRef]
18. *Codice dei Contratti Pubblici*, D. Lgs. April 18, 2016, N. 50. Gazzetta Ufficiale n. 91 April 19, 2016, s.o. n.10. Art. 93, Subsection 1 (Bid Bonds); art. 103, Subsections 1–3, 5 (Performance Bonds). Available online: https://www.codicecontrattipubblici.com/ (accessed on 1 September 2021).

19. *Codice Civile* (Regio Decreto 16 Marzo 1942, n. 262), Art. 1916. Available online: https://www.gazzettaufficiale.it/anteprima/codici/codiceCivile (accessed on 1 September 2021).
20. Directive 2009/138/EC of the European Parliament and of the Council of 25 November 2009. Available online: https://eur-lex.europa.eu/eli/dir/2009/138/oj (accessed on 1 September 2021).
21. Commission Delegated Regulation (EU) 2015/35 of 10 October 2014 Supplementing Directive 2009/138/EC of the European Parliament and of the Council on the Taking-Up and Pursuit of the Business of Insurance and Reinsurance. Available online: https://eur-lex.europa.eu/eli/reg_del/2015/35/oj (accessed on 1 September 2021).
22. Commission Delegated Regulation (EU) 2019/981 of 8 March 2019 Amending Delegated Regulation (EU) 2015/35 Supplementing Directive 2009/138/EC of the European Parliament and of the Council on the Taking-Up and Pursuit of the Business of Insurance and Reinsurance (Solvency II). Available online: https://eur-lex.europa.eu/eli/reg_del/2019/981/oj (accessed on 1 September 2021).
23. Sandstrom, A. *Handbook of Solvency for Actuaries and Risk Managers: Theory and Practice*; Chapman&Hall/CRC Finance, Taylor&Francis Group: London, UK, 2011; ISBN 13-978-1439821305.
24. Olivieri, A.; Pitacco, E. *Introduction to Insurance Mathematics: Technical and Financial Features of Risk Transfers*; Springer: Berlin/Heidelberg, Germany, 2011; ISBN 978-3-642-16028-8.
25. IVASS Regulation No. 38 of 3 July 2018; Article 5, Paragraph 2, Letter E. Available online: https://www.ivass.it/normativa/nazionale/secondaria-ivass/regolamenti/2018/n38/ (accessed on 1 September 2021).
26. Paulusch, J. The Solvency II Standard Formula, Linear Geometry, and Diversification. *J. Risk Financ. Manag.* **2017**, *10*, 11. [CrossRef]
27. Baione, F.; De Angelis, P.; Granito, I. On a capital allocation principle coherent with the Solvency 2 Standard Formula. In Proceedings of the IVASS Conference on Insurance Research, Rome, Italy, 13 July 2017. Available online: https://www.ivass.it/pubblicazioni-e-statistiche/pubblicazioni/att-sem-conv/2017/conf-131407/ (accessed on 1 September 2021).
28. Non-Performing Loans (NPLs) in Italy's Banking System. 2017. Available online: https://www.bancaditalia.it/media/views/2017/npl/ (accessed on 1 August 2021).
29. The Count of Newly Non-Performing Borrowers during the Period Series is Labelled as *TRI30529_351121441*. The Count of Performing Borrowers at the Beginning of the Period Series is Labelled as *TRI30529_351122141*. Data Are Filtered by the ATECO 2007 Economic Sector "*F-Constructors*". Bank of Italy Statistical Database. Available online: https://infostat.bancaditalia.it/inquiry/ (accessed on 1 August 2021).
30. Credit Suisse First Boston. *CreditRisk$^+$, A Credit Risk Management Framework*; Credit Suisse First Boston: London, UK, 1998.
31. Wilde, T. CreditRisk$^+$. In *Encyclopedia of Quantitative Finance*; John Wiley & Sons: Hoboken, NJ, USA, 2010.
32. Giacomelli, J.; Passalacqua, L. Calibrating the CreditRisk$^+$ Model at Different Time Scales and in Presence of Temporal Autocorrelation. *Mathematics* **2021**, *9*, 1679. [CrossRef]
33. Giacomelli, J.; Passalacqua, L. Improved precision in calibrating CreditRisk$^+$ model for Credit Insurance applications. In *Mathematical and Statistical Methods for Actuarial Sciences and Finance-eMAF 2020*; Springer International Publishing: Berlin/Heidelberg, Germany, 2021. [CrossRef]
34. Vandendorpe, A.; Ho, N.D.; Vanduffel, S.; Van Dooren, P. On the parameterization of the CreditRisk$^+$ model for estimating credit portfolio risk. *Insur. Math. Econ.* **2008**, *42*, 736–745. [CrossRef]
35. Higham, N. Computing the nearest correlation matrix—A problem from finance. *IMA J. Numer. Anal.* **2002**, *22*, 329–343. [CrossRef]
36. Krishna, V. *Auction Theory*; Academic Press: San Diego, CA, USA, 2002; ISBN 978-0-12-426297-3.
37. Milgrom, P. *Putting Auction Theory to Work*; Cambridge University Press: Cambridge, UK, 2004; ISBN 978-0-521-55184-7.
38. Fisher, R.A.; Tippett, L.H.C. Limiting forms of the frequency distribution of the largest and smallest member of a sample *Proc. Camb. Phil. Soc.* **1928**, *24*, 180–190. [CrossRef]
39. Gnedenko, B.V. Sur la distribution limite du terme maximum d'une serie aleatoire. *Ann. Math.* **1943**, *44*, 423–453. [CrossRef]
40. Suzuki, A. Investigating Pure Bundling in Japan's Electricity Procurement Auctions. *Mathematics* **2021**, *9*, 1622. [CrossRef]

mathematics

Article

A Fusion Framework for Forecasting Financial Market Direction Using Enhanced Ensemble Models and Technical Indicators

Dushmanta Kumar Padhi [1], Neelamadhab Padhy [1], Akash Kumar Bhoi [2], Jana Shafi [3] and Muhammad Fazal Ijaz [4,*]

[1] Department of Computer Science and Engineering, School of Engineering and Technology, GIET University, Gunupur 765022, Odisha, India; dushmantapadhi@hotmail.com (D.K.P.); dr.neelamadhab@gmail.com (N.P.)
[2] Department of Computer Science and Engineering, Sikkim Manipal Institute of Technology, Sikkim Manipal University, Majitar 737136, Sikkim, India; akashkrbhoi@gmail.com
[3] Department of Computer Science, College of Arts and Science, Prince Sattam bin Abdul Aziz University, Wadi Ad-Dwasir 11991, Saudi Arabia; j.jana@psau.edu.sa
[4] Department of Intelligent Mechatronics Engineering, Sejong University, Seoul 05006, Korea
* Correspondence: fazal@sejong.ac.kr

Abstract: People continuously hunt for a precise and productive strategy to control the stock exchange because the monetary trade is recognised for its unbelievably different character and unpredictability. Even a minor gain in predicting performance will be extremely profitable and significant. Our novel study implemented six boosting techniques, i.e., XGBoost, AdaBoost, Gradient Boosting, LightGBM, CatBoost, and Histogram-based Gradient Boosting, and these boosting techniques were hybridised using a stacking framework to find out the direction of the stock market. Five different stock datasets were selected from four different countries and were used for our experiment. We used two-way overfitting protection during our model building process, i.e., dynamic reduction technique and cross-validation technique. For model evaluation purposes, we used the performance metrics, i.e., accuracy, ROC curve (AUC), F-score, precision, and recall. The aim of our study was to propose and select a predictive model whose training and testing accuracy difference was minimal in all stocks. The findings revealed that the meta-classifier Meta-LightGBM had training and testing accuracy differences that were very low among all stocks. As a result, a proper model selection might allow investors the freedom to invest in a certain stock in order to successfully control risk and create short-term, sustainable profits.

Keywords: stock exchange; stock market; ensemble; cross-validation; LDA; hist gradient boosting; securities exchange; CatBoost

JEL Classification: G10; G15; G17; E17; E44; D53

1. Introduction

Forecasting future stock values has long been a contentious academic issue. For a significant stretch of time, it was assumed that fluctuations in stock values could not be predicted. The share value index is an important part of the financial system since it represents global economic success. Real-world businesses must be watchful of their security as well as their growth. At almost the same moment, investors and analysts were interested in learning about the overall capital market patterns and trends. As a result, correctness in forecasting is critical for stakeholders' well-being. In the midst of the messy and volatile character of stock markets, forecasting future price movements is a difficult topic on which academicians are seeking to improve forecasting models.

Stock value trend forecasting is a masterpiece and fascinating subject drawn by various trained professionals and researchers from fields such as financial engineering, economics, operations research, statistics, and artificial intelligence. Although a large amount of effort has been put in over the last few years, the exact figure of the stock cost and its directions are still challenging to achieve at this point, even though some high-level AI strategies are used. Globalisation of the economy constantly requires developments in the field of computational science and data innovation. In recent years, monetary exercises have been progressively developing in number with the fast financial turn of events, and their varied pattern has additionally become gradually more intricate. The securities exchange assumes a fundamental part of the monetary space of any country [1].

As of now, with the fast improvement of AI and manufactured reasoning in the previous 10 years, an ever-increasing number of market analysts have begun to execute the index value estimating of gaugeable models, have exclusive requirements, and have attempted different strategies [2]. The best standard for deciding on the presentation of the model is to look at the anticipated effects of the model with genuine information. By investigating the current exploration, we can find that even though it is hard to anticipate the securities exchange law precisely, it can foresee the future pattern of the financial exchange somewhat and decrease the dangers looked at by financial backers [3].

Currently, although there are numerous approaches to anticipate the cost of the financial exchange, to all the more likely dissect and manage the information available, as we may found delightful many issues and curtailments, it has furthermore become the point of convergence of examination to come into more critical data. Since traditional insightful techniques have apparent imperfections in taking care of non-linear issues, some machine learning algorithms are brought into securities exchange investigation [4]. A predictive model that can gauge the direction of a stock value development assists financial backers with settling on suitable choices, improve productivity, and consequently decline potential misfortunes. As a result, precise forecast and investigation of the stock market become more challenging and advantageous. For stock value prediction, we must constantly increase determining methods. Previously numerous researchers at local and overseas committed themselves to develop gaugeable monetary frameworks to anticipate index value development. Before the beginning of proficient AI calculations, analysts regularly utilised diverse statistical techniques to fabricate expectation models. There are linear models and non-linear models used for stock price prediction. Most of the linear models come under statistical methods, whereas non-linear models are based on ML algorithms. Implementing a fiscal framework to accomplish precise index value gauging has become a hypothetical and pragmatic work [5]. In principle, the conventional fiscal frameworks and the arising computerised reasoning model can accomplish the expectation of stock costs, yet the forecast impact is very extraordinary [6].

Discovering frameworks with better prescient impacts through model blend and examination is beneficial for some researchers, and it likewise has significant hypothetical importance [7]. In actuality, realised information can be infused into the monetary frameworks to anticipate future information. For example, if the stock value gauge is higher than the end cost of the day, the model predicts that the future stock cost may rise, and financial backers can decide to keep holding the stock to acquire higher venture pay [8].

On the off chance that the stock value gauge is lower than the day's end value, it demonstrates that the stock cost may fall later on. Subsequently, it is incredibly viable to develop a monetary model to acknowledge stock value gauging [9]. Moreover, if you can figure out how to precisely anticipate stock value developments and unpredictability patterns, at that point, it has a significant incentive for nations, recorded organisations, and individual financial backers [10]. As of late, there have been a developing number of studies taking a direction at the course or pattern of developments of financial markets. Now the study gradually increases by looking at the demand and trend of stock markets. Technical investigation and fundamental investigation are two different strategies by which we can foresee the securities exchange. The fundamental investigation relies on precise information on the other variables that affect the securities exchange such as miniature financial aspects; large-scale financial matters; and political and, surprisingly, mental components.

In any case, the information is typically not promptly accessible. The technical examination endeavours to make expectations dependent on past designs. In any case, these models are not, for the most part, evident because of the upheaval [11]. For customary measurable strategies, it is tough to catch the abnormality. In these conventional frameworks, we need to accept a practical connection between information and yield and attempt to fit the data according to that relationship. This has empowered scholastic scientists and business professionals to grow more unsurprising estimation frameworks. Numerous different technologies and strategies have been proposed to embrace and anticipate stock costs through multiple approaches. Yet, the appropriate blend of feature selection and the dynamic behavior of the stock market consistently are an open challenge for researchers to discover a solution. With the expanding accessibility of high-recurrence trading data and the irregularity given by prior models, it is consistently available for local and international researchers to foster a model, which will provide a reliable outcome [12].

1.1. The Inspiration Is as per the Following

Forecasting of the stock market always is an interesting and open challenge problem for researchers [8]. As day to day more information is opening up, we face new difficulties in securing and handling the information to extricate the knowledge and examine the impact on stock costs. Finding the best possible approach for predicting the daily return direction of the financial market is always a challenging and debatable topic [13]. However, the desired goal of this study is to forecast the future market. The most common and fascinating part of this research area in the forecasting of the stock market is its self-sabotaging behaviour. The rapid development of machine learning models tools and technologies always provides opportunities for the researcher to find the hidden truths of the market and analyse the market in their own ways [14]. Identification of proper feature selection increase the performance of prediction of machine learning models [15]. Only a few studies have attempted to identify significant input features [16]. More research is needed on technical indicators for finding an optimal combination of input features for predicting stock prices [17]. However, the performance of forecasting models depends on quality features, and inappropriate feature selection leads to degrading the performance of the model and returns as a biased result. As proper feature selection takes an important role in the model building process in that building a reliable forecasting model which can identify risk factors and provide the positive and negative direction of the market is equally important. Thus, proper selection of algorithms during the model building process is a large challenge for researchers. Past researchers have attempted to adopt hybridisation techniques using either base-level machine learning models or deep learning models, but there is still a question mark as to whether we can hybridise ensemble models; ensemble models are one type of hybridisation of weak learners [18]. Hybridisation of ensemble models can provide better accuracy using voting/averaging techniques. Most researchers' basic selection criteria for finding the best model is to look at the testing accuracy, which is a common and straightforward process that leads to overfitting.

Along these lines, our objective was to develop an ensemble-based hybrid model that learns from the past stock market data and gauges the directional movement of the stock.

1.2. Our Research Contributions in a Nutshell

- We propose a novel framework where six ensemble models are hybridised, minimising the model risk and increasing accuracy.
- A new set of input features were designed, providing a real test for future researchers to think of that combination.

- In this approach, we adopted two-phase overfitting protection. The first is LDA, and the second is the K-fold cross-validation. These techniques were merged into a single framework, making our model a unique one.
- Our model selection process is somewhat different and uncommon. Instead of selecting the model which provides the highest accuracy, we selected the model whose training and testing accuracy difference is minimal, which is very much uncommon and innovative, and this selection process produces a model neither overfitted nor underfitted.
- Specifically, a long time period of data was collected for our experimental setup, which explores the performance level of volatility–stress periods and smooth trending periods and it also examines the persistence of financial crisis and clustering.

The rest of the article is figured out as given below. Section 2 portrays the related work, while Section 3 depicts the Materials and Methods, and in Section 4, we explore our proposed framework. In Section 5, we focus on the exploratory outcomes and discuss critical discoveries in our examination. Finally, in Section 6, we discuss the conclusion part of our paper and the future scope of our study.

2. Related Work

For a long period of time, financial backers and researchers were of the belief that stock cost cannot be anticipated. This conviction appeared due to the efficient market hypothesis (EMH) term coined by Fama [19]. According to Fama, due to the dynamic behaviour and non-stationary nature of financial market data, the financial market cannot be predictable [19]. The EMH says that once a piece of new information is entered into the investment securities, the market reacts instantaneously. Thus, it is impossible to crack the market.

On the other hand, the hypothesis again revised by the hypothesiser and their revised version classified the study into three forms as strong form, semi-strong form, and weak form [20]. The weak form of the hypothesis surmises that using historical prices future stock price cannot be forecasted. The semi-strong form surmises that the stock market behaves instantly as any new information (publicly available) is entered, practically showing there is no opportunity to forecast the market. The third form is the strong form, which deals with both public and private information, which implies that it does not provide financial backers an edge on the lookout. However, some researchers accept the EMH theory, and some researchers have disputed the efficient market hypothesis, both empirically and theoretically [21–24]. According to Nti et al. [25], the amenable involved in the EMH is open for discussion to choose which one is correct. According to Shiller [26], a new era opened for the financial market in the 1990s when behavioural finance was focused on academics. The Nobel Laureate Robert Shiller's [26] investigation revealed that during the period of 1989 to 2000, the up and down of the stock market was influenced by sentiment. At the turn of the century, Thaler [27] implemented behavioural finance to forecast the cave-in of the internet stock boom and accused the broadly held EMH that acknowledged all financial backers as being normal and making fair-minded figures about what is to come. According to Shiller [26], the behavioural finance remains on the opposite side of EMH and puts an idea that the market made changes inside stock always reflects genuine information. Shiller [26] showed that stock costs are very unstable over a short period of time; however, to some degree, the stock market can be predictable over long periods.

Thus, we consider the above authors' outcomes that in the current scenario, there is a chance of prediction of the stock market.

The financial market forecasting is commonly based on two factors: fundamental and technical factors [12,25,28–30]. The fundamental analysis utilises the monetary remaining of the firm, workers, the directorate, management decision policies, monetary status, company's yearly report, asset report, pay reports, earthbound and climatic conditions such as unnatural or catastrophic events, and political information to anticipate the future of the stock market [31–34]. The fundamental factor normally deals with the companies' GDP, CPI, and P/E ratios [35]. For the prediction of stock market, using a fundamental approach is more suitable for long run instead of short run forecasting [36]. The specialised investigators attempt to foresee the securities exchange through the learning of graphs that depict the historical market costs and technical indicators [37–39]. Technical indicators are statistical techniques that are calculated with the help of mathematical formulas using historical prices [40]. The development of artificial intelligence techniques and the increased number of datasets that are easily publicly available brings about new opportunities for researchers to explore something new from the market. According to Tshilidzi [41], the rapid development of AI techniques influences the EMH theory and provides an efficient way to learn from the market. A growing amount of research has been conducted [42–48], finding that post attestation demonstrates that the financial market may be anticipated to some extent [37,49]. Thus, there is a scope for investors to minimise the loss and maximise the profit when dealing with the stock market [50]. In recent studies, the financial market analysis and forecasting basically falls into two categories, i.e., statistical and machine learning [51].

2.1. Statistical Technique

Before the implementation of machine learning techniques, statistical techniques are used to learn the patterns of the stock are and given an approach to dissect and anticipate stocks. A group of statistical approaches are used, i.e., ARIMA, ARMA, GARCH, STAR, EMA, LDA, QDA, and regression techniques for the analysis of the financial market [52], with the ARIMA, EMA, and regression approaches having a predictable capability to some extent [53,54]. As the stock market is dynamic and non-linear in nature, the traditional statistical techniques have suffered a large amount to learn non-linear behaviour, and therefore the emerging machine learning techniques can avoid the limitations of traditional statistical techniques [55].

2.2. Machine Learning Technique

For the forecasting of the stock market, a large number of machine learning algorithms have been implemented [13,16,49,52,56–62]. As is known from previous studies, to predict stock market directional movement using machine learning techniques normally produces better outcomes than any other techniques [63]. Leung et al. [64] found that the exact estimates of the stock worth list development are critical for building effective trading methods such as financial backers that can fence against the expected dangers from the securities exchange. And even though a small amount of improvement on accuracy, its anticipating execution will be profoundly beneficial. Machine learning technique commonly uses two approaches to predict the stock market (a) using a single model to predict the stock market, and (b) using an ensemble of machine learning models [13,60,63,65,66]. The use of ensemble models reported by some researchers found that ensemble models provide better performance than a single predictive model [40,67,68]. According to Fatih et al. [69], there still is little research that has been done to predict the stock market using ensemble models.

As compared to traditional models, machine learning models behave more flexible. There are so many machine learning algorithms that have been applied in previous studies [70]. Examples are logistic regression, support vector machine, k-nearest neighbours, random forest, decision tree [40], and neural networks [49,52,60,71]. As seen in the literature, the most commonly used algorithms for stock market forecasting are support vector machines and artificial neural networks [72]. Milosevic et al. [73] proposed a classification framework to predict the financial market in a long window. They suggest that if the stock value rises 10% in a financial year, we can consider the stock as a good stock; otherwise, it is lousy stock. During their model building process, they extracted 11 fundamental ratios applied to different algorithms as input features. Their study revealed that in differentiation with naïve Bayes and SVM, random forest shows a good F-fcore, i.e., 0.751. Ballings et al. [37] discussed how various ML models have been created for discovering the direction of the stock market. Their study adopted different ensemble machine learning algorithms such as random forest, AdaBoost, neural network, logistic regression, SVR, and KNN, as well as the datasets chosen from European Companies. Their model attempted to predict the price movement of the long-term stock market, and their study revealed that the random forest algorithm performs well in their dataset. Choudhury et al. [74] proposed an ANN model wherein their model used a backpropagation algorithm for the training phase and a multilayer feed-forward network for the testing phase for forecasting the value of a share. Their proposed model provided 0.996 as the regression value. Boonpeng et al. [75] proposed a multi-class classification problem in which their model can classify whether to buy, hold, or sell the stock. The author developed two models, one-against-all and one-against-one neural networks, and compared their performances with the traditional neural network. They concluded that one-against-all neural networks performed better than one-against-one and traditional neural network models, with an accuracy of 72.50%. According to Yang et al. [76], for a successful forecasting model, it is necessary to learn the non-linear factors of a stock. The authors proposed a radial basis function based on SVM with a genetic algorithm that is used for the forecasting of the stock market for the short run.

Dey et al. [77] proposed a model to forecast the stock whose input features are technical indicators. Their model was developed by using XGBoost algorithm, obtaining an accuracy level of 87.99% on the dataset Apple and Yahoo indexes. They compare their proposed model with SVM and ANN and finally revealed that their XGBoost models were the best among them. Basak et al. [19] proposed a framework for classification problems that forecast the price of the stock will increase or decrease. They used random forest and XGBoost classifier, and their study revealed ensemble models perform better if the proper combination of technical indicators is used as input features for a model. According to Ernest et al. [69], the ensemble machine learning models provide superior results in comparison with any individual machine learning model. In their study, they focused on the tree-based ensemble models, and their models were trained with three different stock exchange datasets. Their findings show that the extra trees ensemble classifier performed better than other tree-based classifiers.

Yang et al. [78] proposed an ensemble-based multi-layer feedforward network for the forecasting of the Chinese stock market. Their model was trained with backpropagation and Adam algorithms, whereas an ensemble was created with the help of the bagging approach. The performance of the model may increase if the used dataset is normalised further. Fatih et al. [69] proposed two models using multilayer perceptron with genetic algorithm and particle swarm optimisation. To train their model, they used nine technical indicators, which were recorded as the RMSE of 0.732583 and 0.733063 for MLP–GA and MLP–PSO, respectively. Finally, they concluded that hybrid machine learning methods can improve the accuracy level during forecasting. Wang et al. [79] developed a hybrid model with a combined BPNN, ARIMA, and ESM effort to predict the stock market weekly. The datasets used here were Shenzhen Integrated Index and DJIA. After successful modelling, they attempted to analyse every single framework with the combined ensemble framework. They found hybrid models performed better than traditional individual models and found 70.16% accuracy when forecasting the stock market's direction. Chenglin et al. [80] proposed a model that can accurately predict stock prices' direction. They used a combined model, a combination of SVM and ARIMA, and concluded that the performance of combined models performed better than a single predictive model. Tiwari et al. [81] proposed an ensemble model combined with the Markov framework with a decision tree to forecast the Bombay stock exchange. The proposed model provides an accuracy level of 92.1%, and it was concluded that combined models provide better accuracy than any individual models. A comparative study conducted by Prasad et al. [51] used three different algorithms, namely, XGBoost, Kalman filters, and ARIMA, and two different datasets taken, namely, NSE and NYSE. Their study was based on individual algorithm forecasting capability as well as a hybrid model also developed by them using Kalman filters and XGBoost. Finally, they compared four models and found the ARIMA and XGBoost to show promising results on both datasets, whereas the accuracy of the Kalman filter was not consistent in both datasets. A total of 87.64% accuracy level was maintained by the ARIMA model on the NSE dataset, whereas 79.44% was maintained on NYSE. On the other hand, an 88.66% accuracy level was maintained by the XGBoost model on the NSE dataset, whereas 79.44% was maintained on NYSE. The Kalman filter model showed a promising accuracy level of 89.09% on the NSE dataset, whereas 64.96% was shown on the NYSE dataset. The hybrid model provided 76.79% on the NSE dataset and 70.91% on NYSE dataset. Instead of suggesting the best model among them all, they left the decision for the users in terms of finding the best one among the four. Qiu et al. [66] proposed a combined model LSTM with attention mechanism, i.e., WLSTM+Attention, on three different indexes, finding that the proposed model MSE was less than 0.05. Moreover, they suggested that proper selection of features can improve the predictive capability of the model.

Several authors used technical indicators as input features to train their model. Weng et al. [18] concluded that only by using macroeconomic indicators with a machine learning approach can one predict the stock index efficiently. Markovic et al. [82] implemented different technical indicators in the LS-SVM model to find the trend movement of the stock market in the Southeast European Market, finding technical indicators to have a certain level of prediction power. Valavanis et al. [83] found that approximately 20% of the stock market predictive models used technical indicators as their input features. According to the authors, technical indicators are generally used to learn the flow of complex patterns from specific stock data and to forecast the upcoming behaviours. Fernández et al. [84] found an optimal combination of technical indicators; the researchers developed several indicators suitable for their use and then found an appropriate mix for their predictive model. Andrade et al. [85], discussed how for finding the optimal combination of indicators for modelling, a large amount of effort has been made. However, still, there is no sophisticated, easy technique available for developers to select appropriate technical indicators.

We refer to a summary of various recent research on short-term finance market prediction for more comprehensive and extensive assessments in Table 1.

Table 1. Some of the resent studies based on short-term stock prediction.

Sl.No.	Authors (Year)/Publisher	Dataset Used	Target Output	Forecasting Period	Method Adopted/Model Proposed	Outcome (According to Authors)
1	[86] Nguyen et al. (2019), MDPI	KOSPI 200, S&P 500	Market direction	Short term	Transfer learning + LSTM	According to the author, their proposed model provides satisfactory prediction performance in comparison to SVM, RF, and KNN.
2	[18] Weng et al. (2018), Elsevier	13 U.S. sector indices.	Stock price	One month ahead price	QRF, QRNN, BAGReg, BOOSTReg, ARIMA, GARCH, Deep LSTM	Combination of technical indicators and Ensemble approach can significantly improve the forecasting performance.
3	[87] Bisoi et al. (2018), Elsevier	BSE S&P 500, HIS, FTSE 100	Stock price and stock direction	Short-term	VMD-RKELM	Their proposed model was superior to the SVM, ANN, naïve Bayes, and ARMA.
4	[88] Naik et al. (2019), Springer	NSE data	Market direction (up or down)	Short term (intraday)	Five-layer DNN	Five-layer DNN outperformed 8 to 11% than three-layer ANN.
5	[89] Lahmiri et al. (2019), Elsevier	Six American stocks	Stock price	Short term (intraday)	VMD, PSO, BPNN	According to the authors, their proposed VMD–PSO–BPNN model is superior to the PSO-BPNN model.
6	[12] Dhanya and Yadav (2019), Springer	INDIA, Nifty Index	Stock price	Short-term	ANN, SVR, EMD, EEMD, CEEMDAN	The proposed CEEMDAN-SVR model performed well as compared to other models. The authors want to improve their model using boosting and bagging algorithms and deep-learning algorithms.
7	[66] Jiayu et al. (2020), PLoS ONE	S&P 500, DJIA, HSI	Stock price	Short-term	Attention-based LSTM	The author addressed that their proposed model shows promising results in comparison to widely used LSTM and GRU models.
8	[13] Xiao and David (2019), Springer	US SPDR S&P 500 ETF (SPY)	Stock direction (up or down)	Daily return direction	DNN + PCA	The author revealed the addition of a proper quantity of hidden layers can achieve the highest accuracy, as is also the case in future.
9	[15] Shen et al. (2020), Springer	Chinese stock market	Stock trend	Short-term	PCA + LSTM with feature engineering	The authors stated that instead of focusing only on the best model, to find the best predictive model, it is essential to add an innovative feature engineering approach to improve model performance.
10	[90] Cervelló-Royo et al. (2015), Elsevier	Athens Stock Exchange	Market trend (bull/bear-flag)	Short-term	Fuzzy model + technical indicators	According to the authors, their model is a conservative model wherein during the bull market, small losses and small gain will be found.

The summary given in Table 1 indicates that there is some scientific study in place to predict the stock market in the short run, and the findings are remarkable [28,63,69,91] to some extent, inducing us to take our direction of study in the short run to forecast the stock market. Zulkernine et al. [92] suggested that it is easier to predict the stock market on a long-term basis than via daily stock forecasting. This is because the daily stock forecasting data constantly fluctuate and are full of noise. Thus, in our study based on fusion ensemble models and technical indicators, a brief review was conducted to some extent in our literature survey. According to Weng et al. [18], the prediction results may increase using the voting or averaging technique of different ensemble models. Singh et al. [93] provide their opinion that in recent times the machine learning models have shown promising results. Consequently, we aimed to build a hybrid model which learns from the historical prices in form of indicators and gauges the direction of the stock on the next day.

Some exploration leads to using a combination of unique models to produce their ensemble or hybrid models with technical indicators for the directional movement of individual stocks. Some researchers use pre-existing ensemble models such as XGBoost, CatBoost, etc., for finding the trend of the stock, and some of them use deep learning techniques to forecast the stock market. Thus far, we have noticed that more researchers are trying to develop hybrid models, i.e., a combination of individual machine learning models or deep learning models or their variety. Another challenge researchers face in finding a proper mix of input features is a very fuzzy and tedious job. We believe that there is still a gap in the hybridisation of individual ensemble models into a single frame with a unique combination of input features.

Therefore, after a successful review, we proposed a novel approach wherein ensemble models used a stacking framework, and the stacking framework takes the trained ensemble models as the base-level classifiers. Again, the ensemble models are to be used as the ensemble model meta-classifiers. In addition, we developed a unique combination of input features for the prediction of the stock market.

3. Materials and Methods

When we attempted to develop an ensemble framework, six types of ensemble algorithms were used: XGB classifier, AdaBoost Classifier, Gradient boosting, LightGBM, CatBoost, and Hist gradient boosting as base learners. In this section, we briefly describe the aforementioned base classifiers, and finally, we explain our final framework. Before discussing our framework, we first establish what ensemble learning is and how it works.

3.1. Ensemble Learning

A method takes forecasts from numerous ML models or similar forecasting from the same models at various times to make them more precise results. Forecast from a solitary individual model probably will not produce that many exact outcomes, and thus we require fostering a gathering-based AI model whose prescient limit is much higher than a solitary calculation. Figure 1 depicts the common structure of an ensemble model.

Bagging and boosting are two effective techniques used in machine learning to ensemble the models. Here, in our experiment, we used the boosting technique stacking classifier method; the stacking classifier works on a two-stage process. First, it sequentially creates base learners that are used as the input for the final stage. Then, in the final stage, the stacking classifier builds a meta-classifier using the base learners, which are treated as level-1 weak learners.

3.1.1. Gradient Boosting

Gradient boosting is a type of classifier used to merge different slow learners to produce a robust forecasted model [94]. The following principle describes gradient boosting.

1. Initially, develop an error function, and that function is optimised at the time of model building.

2. Iteratively develop weak models for forecasting.
3. Finally, all the weak models are merged a create a robust model with minimising error function.

Our model-building process begins with a comparatively weak learner model (depending upon our dataset). Then, iteratively, a weak learner is converted to a better classifier F_m(x), and that classifier becomes a robust classifier.

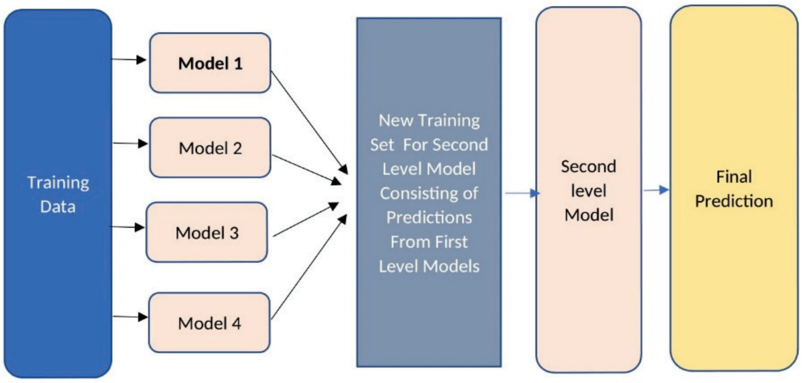

Figure 1. The common structure of an ensemble model [67].

3.1.2. AdaBoost

In 2003, Freund and Robert [95] developed a meta-algorithm and coined its name, AdaBoost; they were awarded a Nobel Prize for their work. According to them, AdaBoost can form an alliance with other machine learning algorithms to increase the accuracy level. Here, in this modelling technique, the weak learners are created sequentially.

During model training, an AdaBoost classifier can be represented as

$$FM_t(z) = \sum_{t=1}^{T} fm_t(z) \tag{1}$$

where the object z is used as an input for each weak learner, i.e., fm_t and a value is returned which represents an object of that class. For instance, in a binary classification problem, the predicted class object and the absolute value depend upon the outcome of the weak learner's sign. Similarly, the sample belongs to the positive class, and the Tth classifier falls into the positive class or otherwise falls into the negative class.

In the training set for each sample, a weak learner generates an output hypothesis $h(z_i)$. During each iteration t, a coefficient αt is assigned to the selected weak learner; for instance, a minimised sum of error terms Em_t will be calculated for the t-stage classifier.

$$Em_t = \sum_{i=1}^{T} Em[Fm_{t-1}(z_i) + \alpha t\, h(z_i)] \tag{2}$$

Here, $Fm_{t-1}(z_i)$ is a robust classifier that is derived from the previous stage training, $Em(Fm)$ denotes the sum of error terms, and $fm_t(z) = \alpha_t h\,(z)$ is the weak classifier which is considered for addition to the final classifier.

3.1.3. Extreme Gradient Boosting (XGBoost)

The XGBoost model is an early experimental model of a Ph.D. student at Washington University [95]. XGBoost is an improved version of the gradient boosting algorithm, which is more scalable and efficient. The features that make the XGBoost algorithm something different are provided here [63]. The automatic feature extraction can be possible. XGBoost supports the regularisation technique to avoid overfitting and has the capability to learn from non-linear datasets. Moreover, the parallelisation feature makes the XGBoost train with multiple CPU cores. It is one of the tree-based ensemble additive models that are composed of multiple base learners. In general, the XGBoost can be represented as

$$\begin{aligned} &\mathrm{F} = \{\mathrm{m1},\ \mathrm{m2},\ \mathrm{m3},\ \mathrm{m4} \ldots\ldots\ldots\ldots \mathrm{m_n}\} \text{ set of base learners} \\ &\text{Predictive model} = \hat{\mathrm{y}}_{\mathrm{i}} = \textstyle\sum_{\mathrm{t}=1}^{\mathrm{n}} \mathrm{mt(xi)} \end{aligned} \tag{3}$$

where $\hat{\mathrm{y}}_{\mathrm{i}}$ is the final predictive model, which is the combination of all weak learners, and x is the input feature for each weak learner, i.e., m.

From the paper [94], we extracted the objective function for XGBoost as given below:

$$Obj(\theta) = \sum_{i=1}^{m} L(z_i, \hat{z}_i) + \sum_{t=1}^{T} \Omega(f_t) \tag{4}$$

By looking at Equation (4), we see that the objective function has two parts; the first part denotes the loss function, i.e., L denotes the training loss of either logistic or squared loss, and the second part represents the addition of each tree's complexity. z_i is actual value and $\hat{z}_i$ is the predicted value, whereas Ω is the regularisation term, T denotes the total number of trees, and f is the function.

3.1.4. The LightGBM

The LightGBM is the most diversely used boosted model that supports parallel training such as extreme gradient boosting [96]. When dealing with multi-dimensional datasets, the LightGBM works much better than the traditional boosting algorithms or XGBoost. Typically, the boosting algorithms split the tree structure horizontally (i.e., level-wise growth), whereas the LightGBM increases the tree structure vertically. Figure 2 shows the growth of tree division between level-wise and leaf-wise.

(a) **(b)**

Figure 2. Tree division level-wise and leaf-wise.

3.1.5. CatBoost

CatBoost is the first Russian machine learning algorithm, developed in the year 2017 by the researchers of Yandex [97]. It is one of the tree-based boosting algorithms. CatBoost stands for categorical boosting, but it deals with absolute values and with other matters (not only with category features but also with regression problems and automatic feature engineering also possible for data). Therefore, compared to the training time of different gradient boosting algorithms, CatBoost takes less time to train. Generally, in a boosting technique, we follow a standard GBT technique for the construction of decision trees, but CatBoost follows two ways of constructing the tree. One is an ordered technique, and the other is a basic technique.

A random permutation technique is applied in the ordered mode for training that follows n number of supporting models, i.e., M1 … Mn, such that the Mi is trained with the help of the earliest I samples in the permutation. In each iteration, for obtaining the residual of the jth sample, the MJ-1 model is used.

3.1.6. Histogram Gradient Boosting

The GBDT framework takes a longer time to train the model as the number of datasets increases, and sometimes the average accuracy level also decreases. Thus, histogram-based gradient boosting is very much effective when there is a large size of datasets [98]. Moreover, this technique reduces training time without losing the accuracy level. Thus, we can say histogram-based gradient boosting is a technique for training faster decision trees used in the gradient boosting ensemble. The splitting principle of histogram-based gradient boosting is as follows: instead of finding the split points on the sorted feature values, the histogram-based algorithm buckets continuous feature values into discrete bins and uses these bins to construct feature histograms during training. Since the histogram-based algorithm is more efficient in both memory consumption and training speed, we developed our work on its basis.

3.2. Dimensionality Reduction Technique

The dimensionality reduction technique is used to minimise the number of features during the training dataset. Dimensionality reduction techniques are applied to the ML models to avoid overfitting issues. As the number of dimensions decreases, the corresponding training parameters also decrease, making the model more straightforward and indicating the degree of freedom. If the number of parameters increases, then the degree of freedom is also high, leading to overfitting the model. That means our model will perform better when the training dataset is provided, but it may not serve better when we provide a test dataset. It is a data preparation technique that will be applied first on the dataset, but it must be remembered that the dataset must be cleaned and scaled before use [99].

3.3. Evaluation Matrices

To examine our proposed model's performances, we used the performance matrices, i.e., accuracy, ROC curve (AUC), and F-score are used. Thus, we took the blend of matrices rather than a solitary one to check the framework's performance. The performance matrices are given below.

$$Precision = \frac{t_p}{t_p + f_p} \tag{5}$$

$$Recall = \frac{t_p}{t_p + f_n} \tag{6}$$

$$Fscore = 2\frac{Precision * Recall}{Precision + Recall} \tag{7}$$

$$Accuracy = \frac{t_p + t_n}{t_p + t_n + f_p + f_n} \tag{8}$$

where

t_p represents the total number of true positive values;
t_n represents the total number of true negative values;
f_p represents the total number of false positive values;
f_n represents the total number of false negative values.

According to Sokolova et al. [100], the area under curve is an appropriate evaluation matrix for classification problems; when the AUC value increases, the prediction performance of the model also increases.

3.4. Tools and Technologies Used

For the technical analysis, Python environment with Anaconda and Google collab was used for model development. Our total model development procedure was executed with an Intel processor (core-i5-1035G1, 1.19 GHz) with 8 GB of memory and a 64-bit Windows operating system.

3.5. Dataset

For our novel approach, we used five different datasets of four different countries, i.e., S&P 500 Index (S&P 500) USA, Dow Jones Industrial Average (DJIA) HONG KONG, Hang Seng Index (HSI) USA, NIKKEI 225 JAPAN, and DAX PERFORMANCE INDEX UNITED KINGDOM. The data on equity indices are daily. All datasets have the equity indices of each trading day.

The period in which the dataset was taken and tested in our model was from 3 January 2000 to 1 July 2019 for DJIA; for S&P 500 Index, the HSI data were from 2 January 2002 to 1 July 2019; the DAX index data were from 12 December 1987 to 18 August 2021; and the NIKKEI 225 data were from 5 January 1965 to 20 August 2021. The Dow Jones Industrial Average index had 4372 tuples, the S&P 500 index had 4904 tuples, the HSI had 4304 tuples, the DAX index had 8495 tuples, and the NIKKEI 225 index had tuples initially. Preliminarily, the records of the dataset had the existing features, i.e., trading volume, high, close, open, and low with the corresponding trading data. 'Yahoo/Finance' portal is a reliable source to download our dataset.

Data Pre-Processing

In previous related studies, there were no specific rules for selecting related input features to forecast the flow direction of the index. Hence, without hesitation, we can say that each technical feature has its hidden behaviour. Using this covert behaviour, the investors try to analyse the current situation and decide whether to buy or sell, according to Weng et al. [18]. Finally, given their conclusion on the technical indicators that analyse the hidden behaviour of these input indicators, it can forecast the monthly closing price of major U.S. indices. Therefore, we used technical indicators and some other features to predict stock index movement in this research.

Once our raw dataset was received, the data needed to be preprocessed. During data preprocessing, we had to follow the following steps:

(a) Generally, the index extracted from the web portal has some existing features that are open, close, low, high, etc. Now, looking at the dataset, we had to handle the null and missing values.

(b) In the second step, we extracted 23 technical indicators using the preexisting dataset described in the previous point. Apart from the technical indicators, we extracted two more features, i.e., the difference between the open and close price, which reflects the increase and decrease of stock value on that day. Another one is to find out the volatility, i.e., the difference between high and low stock prices.

(c) Label generation: In label generation, we constructed a response predicted variable, the binary feedback variable, i.e., $Z_t \in \{0, 1\}$, for individual trading days of our stock. The feedback variable that will be forecast on the T'th day is calculated as

If $Open_t < Close_t$
Then
$Z_t = 1$
Else
$Z_t = 0$
End If

Z_t is the forecast label labelled as 'TREND' that is used as our predicted variable, Open is the opening price of the index on the day, and $Close_t$ is the closing price of the index on the day. Here, we assume that when the Z_t value returns '1', the stock price will increase, and when the Z_t value returns '0', we consider the stock price to have decreased.

(d) Although we are dealing with technical indicators representing a stock's hidden behaviour, we must find a perfect combination of input features with no multi-collinearity issues. The problems with multi-collinearity from a mathematical viewpoint are that the coefficient gauges themselves will, in general, be untrustworthy and that variable is not measurably critical; because of these disadvantages, we ought to consistently check for multi-collinearity in our dataset. For checking multi-collinearity, we have to create a corelation matrix with the help of the correlation function corr(·) (which is used to find the pairwise correlation of all columns in the dataframe) function. This function creates a matrix with a correlation value with the combination of each variable. Therefore, when we diagonally check the matrix, we will obtain correlation values. By looking at the matrix, we have to remove the features whose values are more than 50. Thus, quickly looking at the matrix, we can easily identify the highly correlated values, which should be released. During this feature selection process, we used 23 technical indicators along with seven standard features. After successful correlation testing, we found only seven input features perfectly combined and ready to train our model.

Finally, we found four technical indicators and two derived features, as well as one pre-existing feature, whose descriptions are given below for our experiment. For our technical analysis purposes, we used (Ta-Lib) library, popularly used by traders and researchers for calculating technical indicators. This library can be downloaded from the www.ta-lib.org (accessed on 1 January 2021) website [101–103]. In this study, we employed technical indicators with other dummy variables as our input features for our ensemble models.

Average Directional Movement Index (ADX)

This indicator is used to find out the movement of price in a positive or negative direction.

$$\text{ADXtd} = ((\text{ADXtd-1} * (\text{n\_day} - 1)) + \text{DXtd})/\text{n\_day}$$

Triple Exponential Average (TRIX)

The goal of TRIX is to find out the change of price percentage between two triple SMEA.

$$\text{Trix} = (\text{ema3n} - \text{ema3n-1})/\text{ema3n-1}$$

where ema3n denotes the previous n period's ema.

Percentage Price Oscillator (PPO)

The PPO can be calculated as the difference between the moving averages on n different lengths, i.e., slow-moving average and fast-moving average, and divided by slow-moving average.

$$\text{PPO} = ((\text{F\_ma} - \text{S\_ma})/\text{S\_ma}) * 100$$

where

- F_ma (fast-moving average for a short period);
- S_ma (slow-moving average for a long period).

Ultimate Oscillator (ULT)

ULT indicates whether our stock is oversold or overbought with the goal that we can produce whether we purchase or sell the stocks.

Open

This is a pre-existing feature that denotes any stock price during the opening of the stock in each day.

Open–close

This is an extraction of one feature that indicates the difference between daily transactions' opening and closing values.

High–low

This is also an extracted feature that finds the volatility of each trading day. It is calculated as the difference between the high price and low price value of that day.

(e) After obtaining the useful features, we divided our database into two parts: 75% of the data reserve for training and 25% of the data to test our predictive model.
(f) Finally, in the data processing step, we implemented the scaling technique to normalise our features, which are to be inputted into our model. The statistical description of our three datasets are provided below in Tables 2–6, in which we exploit each of the features of max, min, mean, and standard deviation values of all three datasets.

Table 2. The DJIA index.

Feature	Max	Min	Mean	Standard Deviation
Open	26,833.470700	6547.009766	14,058.426953	4872.465559
Open–Close	1041.839840	−1020.718750	−3.178378	133.911740
ADX	69.127234	8.407117	23.798045	9.135691
TRX	0.208359	−0.456629	0.021683	0.101179
ULT	84.329733	23.405777	54.456530	10.764429
High–Low	1596.648440	25.150390	160.771935	116.553829
PPO	5.590482	−9.709044	0.157680	1.486415

The Dow Jones Industrial Average index had 4372 tuples.

Table 3. The HSI.

Feature	Max	Min	Mean	Standard Deviation
Open	33,335.480470	8351.589844	20,240.157879	5563.590816
Open–Close	1356.910160	−1441.719730	9.175004	198.295114
ADX	57.009177	8.688415	23.420749	8.575097
TRX	0.512277	−0.716833	0.022766	0.157158
ULT	86.389832	18.160742	52.695513	10.444615
High–Low	2060.559570	0.000000	260.297083	177.965756
PPO	7.019466	−11.705714	0.148573	2.112400

The HSI had 4304 tuples.

Table 4. The S&P 500 index.

Feature	Max	Min	Mean	Standard Deviation
Open	2952.709961	679.280029	1539.047669	5563.590816
Open–Close	104.010010	−104.57983	−0.194698	14.922201
ADX	55.046212	8.137028	22.807377	7.958571
TRX	0.235341	−0.569180	0.014379	0.110814
ULT	87.234912	21.561675	54.523968	10.889013
High–Low	125.219971	2.900025	18.132874	12.235494
PPO	7.019466	−10.804408	0.100939	1.587514

The S&P 500 index has 4904 tuples.

Table 5. The NIKKEI 225 index.

Feature	Max	Min	Mean	Standard Deviation
Open	38,921.648438	1020.48999	12,985.66006	8041.057993
Open–Close	1977.890625	−2676.550781	3.268797	158.894006
ADX	75.0401	5.06951	24.674705	10.25368
TRX	0.441637	−0.679374	0.023049	0.142088
ULT	100	4.24892	53.45467	15.33835
High–Low	4206.3	0.00000	142.392	181.764
PPO	9.277873	−14.0134	0.162396	1.948208

The NIKKEI 225 index had 13927 tuples.

Table 6. The DAX index.

Feature	Max	Min	Mean	Standard Deviation
Open	15,948.15	1211.24	6084.276	3693.168
Open–Close	702.86	−508.54	0.680938	73.83413

Table 6. *Cont.*

Feature	Max	Min	Mean	Standard Deviation
ADX	59.474862	7.034225	22.703418	8.590422
TRX	0.399542	−0.562701	0.031402	0.150613
ULT	94.791195	5.549906	53.478123	12.127739
High–Low	921.060546	0.00000	89.470937	79.640647
PPO	6.215173	−16.024076	0.220511	2.126011

The DAX index had 8495 tuples.

4. Proposed Framework

The objective of our study was to develop a framework in which different ensemble models are combined together to form a single predictive model. In a standard stacking framework, the training database is fitted by the base learners known as first-level classifiers. These base-level classifiers after training are used as the input features for the second-level meta classifiers. However, sometimes in ensemble techniques, the level-1 models show overfitting problems. Thus, we simply introduce the dimensional reduction technique to prepare necessary inputs for the first-level classifiers that may not create any over-fitting issues.

We trained all the base-level classifiers in our fusion-based work and converted them into a forecast model; these are used as input for our stacking framework. Hence, we gathered each base classifier's predictive output in our model development process, which is treated as a new set of data for our final model. Thus, categorically, we divided the model into two phases. In the first phase, XGBoost, Adaboost, GB, LightGBM, CatBoost, and Hamming Gradient Boosting are used as level-1 classifiers. The stacking classifiers will be used as a level-2 classifier which is called a meta-classifier and is used to extract the preprocessed hidden features from the level-1 classifiers and combine the level-1 classifiers to make a strong level-2 classifier. The developed framework and pseudo-code are given in Figure 3 and Algorithm 1.

Algorithm 1. Pseudocode for the proposed framework.

Input: Training data $T_d = [x_k, y_k]_{k=1}^{n}$
Output: A boosted stacking meta classifier model denoted as M
Step 1: Implementation of dimensional reduction technique to prepare a training set for base-level classifiers.
Step 2: Initialise E to 6 (number of base-level classifiers)
    For i < −1 to E do
        Read the baselevel classifier m_i
Prepare a training set for first level classifiers
        For t < −1 to T_d do
Train the model
        m_i (t)
        end for
        $fm_i < -m_i$
    end for
Step 3: Read the level-2 meta classifier $[M]_{j=1}^{6}$
Train the meta-classifier with cross-validation technique
    For j < −1 to E do
            For i < −1 to E do
                Ds = Ds + fm_i
            End for
Create a meta-classifier
            M (D_s)
    End for
Step 4: Return (M)

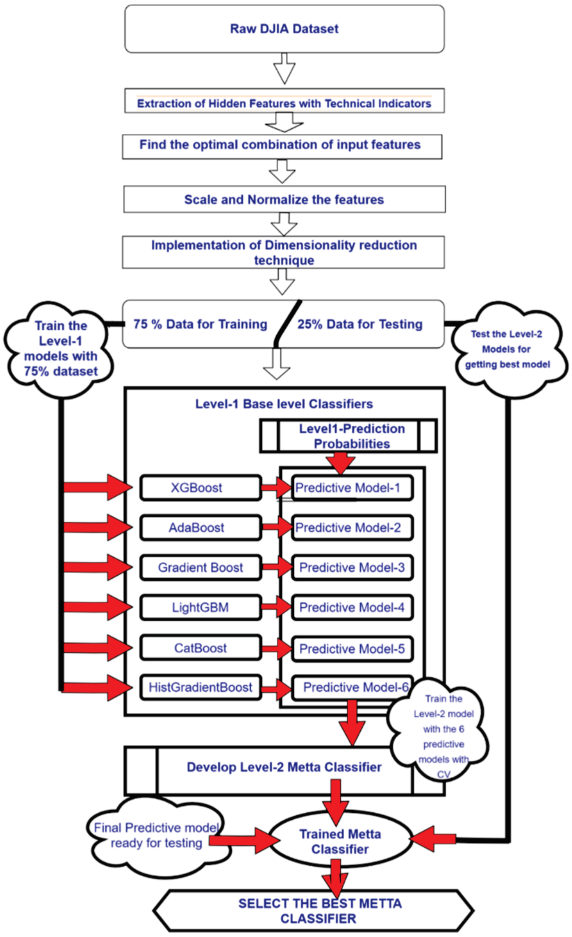

Figure 3. Proposed framework.

We developed a fusion-based ensemble model in this investigation, wherein we used six ensemble algorithms, i.e., XGBoost, Adaboost, LightGBM, GB, CatBoost, and Hamming Gradient Boosting.

Here, we implemented the two most advanced techniques for avoiding overfitting issues during training the models. We took three different datasets and seven input features (detailed description given in Section 3.2). In any model building, process data pre-processing takes an essential role in generating a better predictive model. Thus, in our experiment, we went through the data pre-processing steps; this is elaborated upon in Section 3. After data pre-processing, we considered our models, which were to be trained. During our model building process, we used the stacking framework for the hybridisation of models. In a stacking framework, it takes two steps to develop a hybrid model. Thus, in step 1, it first finds out what base-level classifiers there are; using the base-level classifiers can prepare a training dataset for the second level. Thus, here, we took all our above-mentioned ensemble models as base level classifiers that are labelled as the level-1 classifier.

However, the real challenge came to our notice when we trained the level-1 classifiers with our seven input features. When we trained our base-level classifiers, we simply obtained overfitted models. Thus, to avoid overfitting issues for level-1 models, we implemented the dimensional reduction technique. Different types of dimensional techniques are present; after the successful implementation of this technique, we found the LDA technique was the perfect one for avoiding overfitting in our level-1 models. After implementing the dimensional reduction technique, we successfully developed six ensemble models whose results are given in Tables 7–11.

These base-level classifiers are used as a training dataset for second-level classifiers, a meta-classifier. However, again, another challenge came to our attention in that when we trained the meta classifiers with the help of level-1 classifiers, again, we faced overfitting issues. Thus, to avoid overfitting in our meta-classifier, we implemented the cross-validation technique. Therefore, after the implementation of the cross-validation technique, not only did we avoid overfitting, but we also obtained a generalised model. Tables 13, 15, 17, 19, and 21 and Tables 12, 14, 16, 18, and 20 show the performance level of our meta classifier with cross-validation and without cross-validation, respectively.

5. Results and Discussion

In this segment, we focus on our experimental results during our development process.

5.1. Performance of Base-Level Classifiers

Here, we implemented six ensemble-based boosting classifiers, namely, XGBoost, Adaboost, Gradient Boosting, LightGBM, CatBoost, and HistoGradient Boosting.

These ensemble models were trained with the three datasets, whose accuracy measures are given below in Tables 7–11.

Table 7. The forecast performances of the base level classifiers of DJIA index.

	XGBoost	AdaBoost	GradientBoost	LGBM	CatBoost	Histogr. Boost
Training accuracy	95.33	94.70	94.3	94.27	93.93	94.42
Testing accuracy	93.18	92.99	92.71	92.43	92.99	93.09
Accuracy dif.	2.15	1.71	1.62	1.84	0.94	1.33

Table 8. The forecast performances of the base level classifiers of HSI.

	XGBoost	AdaBoost	GradientBoost	LGBM	CatBoost	Histogr. Boost
Training accuracy	97.24	96.83	96.83	96.48	95.63	96.64
Testing accuracy	95.16	94.97	95.06	95.25	94.21	94.78
Accuracy dif.	2.08	1.86	1.77	1.23	1.42	1.86

Table 9. The forecast performances of the base level classifiers of S&P-500 index.

	XGBoost	AdaBoost	GradientBoost	LGBM	CatBoost	Histogr. Boost
Training accuracy	97.36	96.45	96.09	95.48	95.54	95.04
Testing accuracy	95.01	95.34	94.85	94.43	95.18	94.43
Accuracy dif.	2.35	1.11	1.24	1.05	0.36	0.61

Table 10. The forecast performances of the base level classifiers of DAX PERFORMANCE index.

	XGBoost	AdaBoost	GradientBoost	LGBM	CatBoost	Histogr. Boost
Training accuracy	90.78	90.06	92.06	88.70	89.32	89.40
Testing accuracy	88.94	89.08	88.50	87.98	89.18	89.37
Accuracy dif.	1.84	0.98	3.56	0.72	0.14	0.03

Table 11. The forecast performances of the base level classifiers of NIKKEI 225 index.

	XGBoost	AdaBoost	GradientBoost	LGBM	CatBoost	Histogr. Boost
Training accuracy	91.44	90.54	92.50	90.11	90.63	90.66
Testing accuracy	89.33	89.30	88.69	88.55	89.42	89.30
Accuracy dif.	2.11	1.24	3.81	1.56	1.21	1.36

We remarked that we used six ensemble boosted models whose training accuracy and test accuracy are given by looking at Table 7 above. Hence, the model XGBoost showed a training accuracy of 95.33 and a testing accuracy of 93.18. The model Adaboost showed a training accuracy of 94.70 and testing accuracy of 92.99, and Gradient Boosting showed a training accuracy of 94.33 and testing accuracy of 92.71; LightGBM showed a training accuracy of 94.27 and testing accuracy of 92.43; CatBoost showed a training accuracy of 93.93 and testing accuracy of 92.99; and finally, the HistoGradient Boosting model showed a training accuracy of 94.42 and testing accuracy of 93.09.

Observing the cells of accuracy difference shows the difference in accuracy between the training accuracy set and the testing accuracy set. For our information purpose, it is always a good practice that develops a model as much as a generalised model. A generalised model is a model with little or no difference between the training and testing accuracies. Hence, Table 7 shows a small variation, i.e., 0.94 to 2.15, between training and testing accuracy; the less indifference in the model is more common. Here, our CatBoost and HistoGradient boosting models were more generalised than the other four models. After this, these six models' predictive outputs were used as input features for the meta-classifier.

We remarked that we used six ensemble boosted models whose training accuracy and test accuracy are given in Table 8. The model XGBoost showed a training accuracy of 97.24 and a testing accuracy of 95.16; the model Adaboost showed a training accuracy of 96.83 and a testing accuracy of 94.97; Gradient Boosting showed a training accuracy of 96.83 and testing accuracy of 95.06; LightGBM shows training accuracy 96.48 and testing accuracy 94.21; CatBoost showed a training accuracy of 95.63 and testing accuracy of 94.21; and finally, the HistoGradient Boosting model showed a training accuracy of 96.64 and testing accuracy of 94.78. From here, all these models were ready to train our meta-classifier, which was our primary goal in terms of developing a generalised model.

As shown in Table 9, we found six models were developed whose training and testing accuracy were given. The model XGBoost showed a training accuracy of 97.36 and a testing accuracy of 95.01; the model Adaboost showed a training accuracy 96.45 and a testing accuracy of 95.34; Gradient Boosting showed a training accuracy of 96.09 and testing accuracy of 94.85; LightGBM showed a training accuracy of 95.48 and testing accuracy of 94.43; CatBoost showed a training accuracy of 95.54 and testing accuracy of 95.18; and finally, the HistoGradient Boosting model showed a training accuracy of 95.04 and a testing accuracy of 94.43. From here, all these models were ready to train our meta-classifier, which was our primary goal in terms of developing a generalised model.

We used six ensemble boosted models whose training accuracy and test accuracy are shown in Table 10. The model XGBoost showed a training accuracy of 90.78 and a testing accuracy of 88.94; the model Adaboost showed a training accuracy of 90.06 and a testing accuracy of 89.08; Gradient Boosting showed a training accuracy of 92.06 and a testing accuracy of 88.50; LightGBM showed a training accuracy of 88.70 and a testing accuracy of 87.98; CatBoost showed a training accuracy of 89.32 and a testing accuracy of 89.18; and finally, the HistoGradient Boosting model showed a training accuracy of 89.40 and a testing accuracy of 89.37. Here, the HistoGradient Boosting provided the lower accuracy difference, i.e., 0.03 whereas the Gradient Boost provided the highest accuracy difference, i.e., 3.56.

We used six ensemble boosted models whose training accuracy and test accuracy are given in Table 11. The model XGBoost showed a training accuracy of 91.44 and a testing accuracy of 89.33; the model Adaboost showed a training accuracy of 90.54 and a testing accuracy of 89.30; Gradient Boosting showed a training accuracy of 92.50 and a testing accuracy of 88.69; LightGBM showed a training accuracy of 90.11 and a testing accuracy of 88.55; CatBoost showed a training accuracy of 90.63 and a testing accuracy of 89.42; and finally, the HistoGradient Boosting model showed a training accuracy of 90.66 and a testing accuracy of 89.30. Here, the CatBoost provided the lowest accuracy difference, i.e., 1.21, whereas the Gradient Boost provided the highest accuracy difference, i.e., 3.81.

5.2. Performances of Fusion-Based Meta-Classifiers

In this section, we attempt find out the best combination of meta-classifiers. In the development of the meta-classifiers, we used the stacking cross-validation technique to combine level-1 predictive outputs and finally developed the metaclassifiers. The aim of our experiment was to find out a generalised model. The model was generalised when the training accuracy and testing accuracy were both are close to each other.

5.2.1. Performances of DJIA Index

Table 12 shows meta-classifiers' structure without cross-validation and Table 13 shows the construction of meta-classifiers with cross-validation.

Table 12. Construction of predictive meta-classifiers without cross-validation of DJIA index.

Models	Training Accuracy	Testing Accuracy	Accuracy	Difference
Meta-XGBoost	95.51	93.37	93.37	2.14
Meta-AdaBoost	95.33	93.28	93.28	2.05
Meta-G.B	95.48	93.18	93.18	2.3
Meta-LightGBM	95.45	93.28	93.28	2.19
Meta-CatBoost	95.11	93.28	93.28	1.84
Meta-H.G.boost	95.48	93.37	93.37	2.11

Table 13. Construction of predictive meta-classifiers with cross-validation of DJIA index.

Models	Training Accuracy	Testing Accuracy	Accuracy	Difference
Meta-XGBoost	93.99	92.99	93	1
Meta-AdaBoost	93.33	92.71	92.72	0.62
Meta-GB	94.55	93.09	93.09	1.46
Meta-Light GBM	93.33	93.27	93.28	0.06
Meta-CatBoost	93.96	92.99	93.00	0.97
Meta-H.G.Boost.	93.96	93.46	93.46	0.5

By observing Tables 12 and 13, we can say that the accuracy value of all models in Table 12 without C.V and in Table 13 with C.V seemed to be the same, with a negligible difference. However, when we came to find a generalised model with good predictive power, then Table 13 with cross-validation provides a promising result.

Table 13 found the six meta-classifiers, i.e., XGBoost, AdaBoost, GB, LightGbm, CatBoost, and HistoGradient Boosting. Out of the six meta classifiers, some of them had good accuracy. Some of them were better, but the LightGBM and HistoGradient Boosting provided good, promising results as they tend to prove themselves as generalised models with good predictive power. The model was generalised when the training accuracy and testing accuracy results were nearer to each other with a small difference or no difference. The meta classifier LightGBM provided training accuracy of 93.33 and testing accuracy of 93.27, with a negligent difference, i.e., 0.06. The meta-classifier HistoGradient Boosting provided a training accuracy of 93.96 and a testing accuracy of 93.46 with a difference of 0.5.

When we compared both models, we found that the meta classifier HistGrandBoosting showed the highest predictive accuracy, i.e., 93.46, whereas LightGbm showed 93.27 as its predictive accuracy; this means in terms of comparison of accuracies, HistGrandboosting proved itself to be more promising than LightGbm. However, if we were aiming for a generalised model with reasonable accuracy, then we can say that LightGbm is the perfect one because the training and testing accuracy difference was only 0.06, providing excellent promising power in prediction with both training and testing accuracy, inducing us to consider it as the best model among all.

5.2.2. Performance of HSI

In Table 14, we show the six meta classifiers that were developed without cross-validation technique; the purpose of this development was to bring about a conclusion that instead of using a single model, if we were to use a combination of models for prediction, then definitely our combined approach can predict better than an individual one.

Table 14. Construction of predictive meta-classifiers without cross-validation of HSI.

Models	Training Accuracy	Testing Accuracy	Accuracy	Difference
Meta-XGBoost	97.53	94.97	94.97	2.56
Meta-AdaBoost	96.90	95.16	95.16	1.74
Meta-G.B	97.53	94.97	94.97	2.56
Meta-LightGBM	97.46	94.97	94.97	2.49
Meta-CatBoost	96.71	95.44	95.44	1.27
Meta-H.G.boost	97.53	94.97	94.97	2.56

By looking at Table 1, we can see that the meta-classifiers were more generalised when the cross-validation technique was applied. For example, from this table, the Meta-H.G Boost and Meta-LightGBM showed remarkable performance, whereas the Meta-H.G Boost provided the highest testing accuracy, i.e., 95.35, and Meta-LightGBM provided 94.97 as the testing accuracy. However, by looking at the differences in accuracies, we found that the Meta-H.G Boost provided a difference of 0.31, whereas Meta-LightGBM provided a difference of 0.12, which was the slightest difference among all six meta-classifiers. Thus, the Meta-LightGBM is the more generalised model among all the models, which are shown in Table 15.

Table 15. Construction of predictive meta-classifiers with cross-validation of HSI.

Models	Training Accuracy	Testing Accuracy	Accuracy	Difference
Meta-XGBoost	95.41	94.30	94.30	1.11
Meta-AdaBoost	95.47	94.78	94.78	0.69
Meta-G.B	96.55	95.16	95.16	1.39
Meta-LightGBM	95.09	94.97	94.97	0.12
Meta-CatBoost	96.48	95.25	95.25	1.23
Meta-H.G.boost	95.66	95.35	95.35	0.31

5.2.3. Performance of S&P-500 Index

As shown in Table 16, we found six meta-classifiers that were developed without using the cross-validation technique. From Table 16, we can see that the meta-classifiers Meta-G.B, LightGBM, and CatBoost had testing accuracies of 95.01, and Meta-XGBoost was 94.76; Meta-AdaBoost had a training accuracy of 95.34 and Meta-H.G.boost had a training accuracy of 94.18.

Table 16. Construction of predictive meta-classifiers without cross-validation of S&P-500 index.

Models	Training Accuracy	Testing Accuracy	Accuracy	Difference
Meta-XGBoost	97.53	94.76	94.76	2.77
Meta-AdaBoost	96.87	95.34	95.34	1.53
Meta-G.B	97.47	95.01	95.01	2.46
Meta-LightGBM	97.36	95.01	95.01	2.35
Meta-CatBoost	97.45	95.01	95.01	2.44
Meta-H.G.boost	97.00	94.18	94.18	2.82

As shown in Table 17, we found that the meta-classifiers were more generalised when the cross-validation technique was applied. Table 17 shows that the Meta-AdaBoost, Meta-GB, and Meta-CatBoost showed the highest testing accuracies, i.e., at 95.43, whereas the Meta-LightGBM provided 94.68 as the testing accuracy. However, by looking at the accuracy difference, we found that Meta-XGBoost had a difference of 0.28, whereas Meta-LightGBM was 0.05, which was the slightest difference among all six meta-classifiers. Thus, Meta-LightGBM was the more generalised one among all.

Table 17. Construction of predictive meta-classifiers with cross-validation of S&P-500 index.

Models	Training Accuracy	Testing Accuracy	Accuracy	Difference
Meta-XGBoost	95.54	95.26	95.26	0.28
Meta-AdaBoost	96.17	95.43	95.43	0.74
Meta-G.B	96.81	95.43	95.43	1.38
Meta-LightGBM	94.73	94.68	94.68	0.05
Meta-CatBoost	96.28	95.43	95.43	0.85
Meta-H.G.boost	95.54	95.18	95.18	0.36

5.2.4. Performance of DAX Index

As shown in Table 18, we found six meta-classifiers that were developed without using the cross-validation technique. Here, by looking at the table, we found the training accuracy of all models appeared same with small differences, such as that we observe the testing accuracy, all model results looked the same and only a small variation was found. However, our result finding revealed that the accuracy difference was increased in DAX index, and therefore there was a chance of underfitting of the model.

Table 18. Construction of predictive meta-classifiers without cross-validation of DAX index.

Models	Training Accuracy	Testing Accuracy	Accuracy	Difference
Meta-XGBoost	92.41	88.26	94.76	4.15
Meta-AdaBoost	92.06	88.50	95.34	3.56
Meta-G.B	92.41	88.26	95.01	4.15
Meta-LightGBM	92.37	88.31	95.01	4.06
Meta-CatBoost	92.41	88.26	95.01	4.15
Meta-H.G.boost	92.41	88.26	94.18	4.15

As shown in Table 19, we found that the meta-classifiers were more generalised when the cross-validation technique was applied. In Table 19, by looking at the accuracy difference, we see that Meta-AdaBoost provided the highest difference, i.e., 0.47, whereas Meta-LightGBM provided 0.10, which was the smallest difference among all six meta-classifiers. Thus, the Meta-LightGBM was the more generalised one among all models. When we applied the cross-validation technique in our model, we found the models were more generalised and the training and testing accuracy difference was significantly less.

Table 19. Construction of predictive meta-classifiers with cross-validation of DAX index.

Models	Training Accuracy	Testing Accuracy	Accuracy	Difference
Meta-XGBoost	88.66	88.22	94.30	0.44
Meta-AdaBoost	89.70	89.23	94.78	0.47
Meta-G.B	88.92	88.60	95.16	0.32
Meta-LightGBM	85.10	85	94.97	0.10
Meta-CatBoost	96.47	96.25	95.25	0.22
Meta-H.G.boost	96.14	95.93	95.35	0.21

5.2.5. Performance of NIKKEI 225 Index

As is shown in Table 20, we found six meta-classifiers that were developed without using the cross-validation technique. Here, we found the training accuracy of all models appeared the same with a small difference such that that if we looked at the testing accuracy, all model results appeared the same, with only a small variation found. However, our result findings revealed that the accuracy difference was increased in NIKKEI 225 index when there was no cross-validation technique applied, which may lead to underfitting of a model.

Table 20. Construction of predictive meta-classifiers without cross-validation of NIKKEI 225 index.

Models	Training Accuracy	Testing Accuracy	Accuracy	Difference
Meta-XGBoost	92.65	88.67	94.76	3.98
Meta-AdaBoost	92.53	88.81	95.34	3.72
Meta-G.B	92.65	88.67	95.01	3.98
Meta-LightGBM	92.65	88.67	95.01	3.98
Meta-CatBoost	92.65	88.67	95.01	3.98
Meta-H.G.boost	92.65	88.67	94.18	3.98

As we can see in Table 21, we found that the meta-classifiers were more generalised when the cross-validation technique was applied. By looking at the accuracy difference, we found that the Meta-XGBoost provided the highest difference, i.e., 1.63, whereas Meta-LightGBM provided 0.22, which was the smallest difference among all six meta-classifiers. Thus, Meta-LightGBM was the more generalised one among all models. When we applied cross-validation technique in our model, we found the models were more generalised and the training and testing accuracy differences were significantly less.

Table 21. Construction of predictive meta-classifiers with cross-validation of NIKKEI 225 index.

Models	Training Accuracy	Testing Accuracy	Accuracy	Difference
Meta-XGBoost	90.84	89.21	94.30	1.63
Meta-AdaBoost	90.62	89.62	94.78	1
Meta-G.B	90.34	89.62	95.16	0.72
Meta-LightGBM	88.60	88.38	94.97	0.22
Meta-CatBoost	90.55	89.71	95.25	0.84
Meta-H.G.boost	89.58	88.67	95.35	0.91

5.3. Evalution Matrices of Meta-LightGBM

From the above experiment and discussion, we found the Meta-LightGBM model to be a more generalised model. Thus, in this paragraph, we explain the performance measurement technique of Meta-LightGBM of different datasets used in our experiment. If we are to access the accuracy of a classification model and the quality of prediction, then we must extract the classification report of that model. This classification report provides the results of metrices, i.e., recall, precision, and f1-score on the basis of class.

Table 22 shows the classification report of meta-classifier LightGBM, and the model had an AUC score of 93.36. When our model worked on the DJIA index, the performance of the metrics for class 0 was 0.93, recall was 0.93, and the F1-score was 0.93, whereas for class 1, the precision was 0.94, recall was 0.93, and the F1-score was 0.94. As we found, the AUC score was much closer to 1, which indicates that our model performance was good enough for this dataset.

Table 22. Classification report of Meta-LightGBM for DJIA index.

Precision	Recall	F1-Score
0.93	0.93	0.93
0.94	0.93	0.94

Table 23 shows the classification report of meta-classifier LightGBM, and the model had an AUC score of 95.43. When our model worked on the S&P 500 Index, the performance of the metrics of class 0 was 0.97, recall was 0.94, and the F1-dcore was 0.95, whereas for class 1, the precision was 0.94, recall was 0.97, and the F1-score was 0.95. As we see, the AUC score was much closer to 1, which indicates that our model performance was good enough for this dataset.

Table 23. Classification report of Meta-LightGBM in S&P 500 index.

Precision	Recall	F1-Score
0.97	0.94	0.95
0.94	0.97	0.95

Table 24 shows the classification report of Metaclassifier LightGBM, and the model had an AUC score of 95.26 for HSI. When our model worked on the HSI, the performance of the metrics of class 0 was 0.95, recall was 0.96, and the F1-score was 0.96, whereas for class 1, the precision was 0.96, recall was 0.94, and the F1-score was 0.95. As we see, the AUC score was much closer to 1, which indicates that our model performance was good enough for this dataset.

Table 24. Classification report of Meta-LightGBM for HSI.

Precision	Recall	F1-Score
0.95	0.96	0.96
0.96	0.94	0.95

Table 25 shows the classification report of meta-classifier LightGBM, and the model had an AUC score of 84.71 for DAX index. When our model worked on the DAX index, the precision metrics of class 0 was 0.86, recall was 0.87, and the F1-score was 0.87, whereas for class 1, the precision was 0.83, recall was 0.82, and the F1-score was 0.83. In comparison to the previous three datasets, the AUC scored somewhat less but was still good enough for a model performance.

Table 25. Classification report of Meta-LightGBM for DAX index.

Precision	Recall	F1-Score
0.86	0.87	0.87
0.83	0.82	0.83

Table 26 shows the classification report of meta-classifier LightGBM, and the model had an AUC score of 84.41 for NIKKEI 225 index. When our model worked on the NIKKEI 225 index, the precision metrics of class 0 was 0.90, recall was 0.94 and the F1-score was 0.92, whereas for class 1, the precision was 0.84, recall was 0.75, and the F1-score was 0.79. In comparison to the previous three datasets, the AUC scored somewhat less but was still good enough for a model performance.

Table 26. Classification report of Meta-LightGBM for NIKKEI 225 index.

Precision	Recall	F1-Score
0.90	0.94	0.92
0.84	0.75	0.79

The Figure 4 shows a graphical presentation of AUC scores, and Table 27 shows the AUC scores of the Meta-LightGBM model of different datasets.

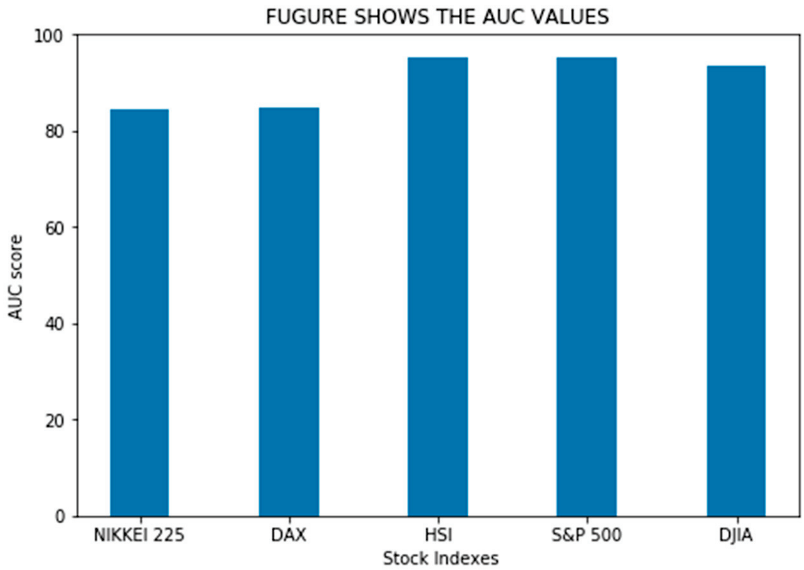

Figure 4. Bar graph shows the AUC score of Meta-LightGBM of all datasets.

Table 27. The AUC scores of all datasets of Meta-LightGBM model.

INDEX	AUC SCORE
NIKKEI 225	84.41
DAX	84.71
HSI	95.26
S&P 500	95.43
DJIA	93.36

5.4. Forecast Accuracy Comparison with Past Work

For a comparative study, we took the suggested model of Qiu et al. [66], i.e., WLSTM+Attention with our proposed model, i.e., Meta-LightGBM. As a benchmark, we took the mean absolute error (MAE) of both the models. The model which gave less MAE value was the best predictive model [40]. As shown in Table 28, our proposed model showed a lower MAE value than the WLSTM+Attention model. Thus, we can say the proposed model can perform better than the WLSTM+Attention model.

Table 28. Comparison of our suggested model with that past work.

Name of the Index	Our Proposed Model (Meta-LightGBM)	WLSTM+Attentation
	MAE	MAE
S&P 500	0.0481	0.1935
HSI	0.0464	0.2453
DJIA	0.0672	0.1569

5.5. Practical Implications

Every investor's 'dream' is to be able to properly anticipate the stock price and, as a result, compute the expected return. The proposed method has the ability to provide investors with useful information. Nowadays, the ML-based tools provide recommendations about specific stocks for the investors so that the investor gains a preliminary idea and can minimise the losses on investment. Artificial intelligence has a genuine effect on monetary exchange by mining significant data and providing modest and effectively accessible apparatuses that advantage everybody, not simply corporates. The speculation choices made by AI will be determined, exact and fair, not at all like those made by humans, who are evidently excessively enthusiastic about the exchange of securities. The proposed model may be used to develop new trading techniques or to manage stock portfolios by changing equities on the basis of trend predictions. It will help different financial institutions to gather information about the movement of the stock so that they can guide their investors to book profits and minimise the losses. Furthermore, it also provides a new direction for future researchers on how ensemble models are hybridised with different combinations of technical indicators and what the outcome will be when different parameters are tuned. Our experimental outcome revealed that the meta-classifier LightGBM had less error differences between training and testing accuracy that made our model more generalised. With the help of our model, any investor can minimise the losses during trading.

6. Conclusions and Future Scope

This study is based on fusion of ensemble models with technical indicators and extracted features to develop an evolutionary ensembled framework for forecasting stock market swings. During the model building process, our goal was to select a generalised model whose training and testing accuracy difference was minimal, instead of finding a model which provides the highest accuracy so that the investor and financial decision-makers can minimise the biased result using this proposed model. For our novel approach, we randomly selected five different indexes of four different countries, using seven features: four technical indicators, two derived features from the pre-existing elements, and the open price of the stock indices. Six ensemble models were used as base classifiers in layer one: XGBoost, LightGBM, AdaBoost, Gradient Boosting, CatBoost, and HistoGradient Boosting. During the individual modelling process (layer one modelling), we used a dynamic reduction algorithm, i.e., LDA, to generate probable input for the next layer classifier to prevent overfitting.

Our fact-finding results revealed that when we fused the ensemble models and developed a meta-classifier without using cross-validation technique, the fusion models training and testing accuracy difference was not good in comparison to the fusion models that were based on cross-validation technique. Sometimes, we found a fusion model provided better performance than a single predictive model, but the fusion models which were trained with cross validation technique showed promising results. During our experiment, we found that when the data size was increased, the performance of the model sometimes decreased. Our goal was to find a fusion model that offered a minimal overfitting and underfitting level, i.e., the training and testing accuracy difference should be significantly less. By looking at Table 29, we can summarise that the meta-classifier Meta-LightGbm is a model that had minimal training and testing accuracy difference in all the indexes. The recorded accuracy differences are shown in Table 29. Instead of only focusing on accuracy if we consider both accuracy and generalised acceptance, the meta-classifier Meta-LightGBM with cross-validation was shown to be more promising than all other predictive models. The run time of the model (Metta-LightGBM) on different datasets depended upon different constraints, such as the hardware configuration of the system in which the model was run though Googlecolab platform and the running time of our proposed model on S&P 500 index was 5 s, DJIA was 4 s, HSI was 4 s, NIKKEI 225 index was 40 s, and DAX index was 14 s in order to train the model. From this, we found that the execution time may differ as the number of datasets increases.

Table 29. Performance summary of Meta-LightGBM of all datasets.

	Training Accuracy	Testing Accuracy	Difference	Dataset
Meta-LightGBM	94.73	94.68	0.05	S&P-500 Index
Meta-LightGBM	95.09	94.97	0.12	HSI Index
Meta-LightGBM	93.33	93.27	0.06	DJIA Index
Meta-LightGBM	85.10	85.00	0.10	DAX Index
Meta-LightGBM	88.60	88.38	0.22	NIKKEI 225 Index

Finally, we can say that the fusion of ensemble models can be more generalised when we apply a cross-validation technique; not only will we improve the predictive accuracy, but also we will obtain a generalised model whose training and testing accuracy are very much closer to each other. Thus, we obtain a model which is neither overfitted nor underfitted.

Limitations and Future Work

Despite our proposed methodology's great predictive performance, there are some constraints which may be worked over in future. Our current study examined only one day ahead in terms of stock direction prediction, and therefore the study must be extended in future for long-term prediction. Our current study focused only on four countries' stock exchanges, but it may be examined and extended to study more stock exchanges of different countries. The research group did not consider any other information sources such as fundamental analysis or sentiment analysis, and thus this dataset must be considered in future experiments. In future work, we will use the factorisation machines and observe how they help in predicting the stock market behaviour [104–107].

Author Contributions: Conceptualisation, D.K.P., N.P., M.F.I., A.K.B. and J.S.; data curation, D.K.P. and N.P.; formal analysis, D.K.P., J.S., A.K.B. and M.F.I.; funding acquisition, M.F.I. and J.S.; investigation, D.K.P., N.P. and A.K.B.; methodology, D.K.P., N.P., M.F.I., A.K.B. and J.S.; project administration, M.F.I., A.K.B. and J.S.; resources, M.F.I., A.K.B. and J.S.; software, D.K.P. and N.P.; supervision, N.P., M.F.I., and A.K.B.; validation D.K.P., N.P. and A.K.B. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by Sejong University Research fund.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Publicly available datasets were analysed in this study. This data can be found on the 'Yahoo/Finance' portal. The data used to support the findings of this study are available from the first author upon request.

Acknowledgments: Jana Shafi would like to thank the Deanship of Scientific Research, Prince Sattam bin Abdul Aziz University, for supporting this work.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

CLAHE	Contrast-limited adaptive histogram equalisation
PSNR	Peak signal-to-noise ratio
PCA	Principal component analysis
ICA	Independent component analysis
FF	Firefly
HO	Hybrid optimisation
MAE	Mean absolute error

References

1. Adam, K.; Marcet, A.; Nicolini, J.P. Stock Market Volatility and Learning. *J. Financ.* **2016**, *71*, 33–82. [CrossRef]
2. Kunze, F.; Spiwoks, M.; Bizer, K. The usefulness of oil price forecasts—Evidence from survey predictions. *Manag. Decis. Econ.* **2018**, *39*, 427–446. [CrossRef]
3. Agustini, W.F.; Affianti, I.R.; Putri, E. Stock price prediction using geometric Brownian motion. *J. Phys. Conf. Ser.* **2018**, *974*, 012047. [CrossRef]
4. Dinh, T.A.; Kwon, Y.K. An empirical study on importance of modeling parameters and trading volume-based features in daily stock trading using neural networks. *IEEE Inform.* **2018**, *5*, 36. [CrossRef]
5. Mehdizadeh, S.; Sales, A.K. A comparative study of autoregressive, autoregressive moving average, gene expression programming and Bayesian networks for estimating monthly streamflow. *Water Resour. Manag.* **2018**, *32*, 3001–3022. [CrossRef]
6. Zhang, Y.; Song, W.; Karimi, M.; Chi, C.H. Kudreyko, AFractional autoregressive integrated moving average and finite-element modal: The forecast of tire vibration trend. *IEEE Access* **2018**, *6*, 40137–40142. [CrossRef]
7. Li, Q.; Cao, G.; Wei, X. Relationship research between meteorological disasters and stock markets based on a multifractal detrending moving average algorithm. *Int. J. Mod. Phys.* **2018**, *32*, 1750267. [CrossRef]
8. Petukhova, T.; Ojkic, D.; McEwen, B.; Deardon, R.; Poljak, Z. Assessment of autoregressive integrated moving average (ARIMA), generalized linear autoregressive moving average (GLARMA), and random forest (RF) time series regression models for predicting influenza a virus frequency in swine in Ontario, Canada. *PLoS ONE* **2018**, *13*, e0198313. [CrossRef]
9. Wang, D.; Liang, Z. Afuzzy set-valued autoregressive moving average model and its applications. *Symmetry* **2018**, *10*, 324. [CrossRef]
10. Rui, R.; Wu, D.D.; Liu, T. Forecasting stock market movement direction using sentiment analysis and support vector machine. *IEEE Syst. J.* **2018**, *13*, 60–770.
11. Nabipour, M.; Nayyeri, P.; Jabani, H.; Mosavi, A.; Salwana, E.S.S. Deep Learning for Stock Market Prediction. *Entropy* **2020**, *22*, 840. [CrossRef]
12. Jothimani, D.; Surendra, S.Y. Stock trading decisions using ensemble based forecasting models: A study of the Indian stock market. *J. Bank. Financ. Technol.* **2019**, *3*, 113–129. [CrossRef]
13. Zhong, X.; Enke, D. Predicting the daily return direction of the stock market using hybrid machine learning algorithms. *Financ. Innov.* **2019**, *5*, 24. [CrossRef]
14. Yu, L.; Hu, L.; Tang, L. Stock Selection with a Novel Sigmoid-Based Mixed Discrete-Continuous Differential Evolution Algorithm. *IEEE Trans. Knowl. Data Eng.* **2016**, *28*, 1891–1904. [CrossRef]
15. Shen, J.; Shafiq, M.O. Short-term stock market price trend prediction using a comprehensive deep learning system. *J. Big Data* **2020**, *7*, 66. [CrossRef] [PubMed]
16. Qiu, M.; Song, Y.; Akagi, F. The case of the Japanese stock market. *Chaos Solitons Fractals* **2016**, *85*, 1–7. [CrossRef]
17. Salim, L. A Technical Analysis Information Fusion Approach for Stock Price Analysis and Modeling. *World Sci. Res.* **2018**, *17*, 1850007.
18. Weng, B.; Martinez, W.; Tsai, Y.-T.; Li, C.; Lu, L.; Barth, J.R.; Megahed, F.M. Macroeconomic indicators alone can predict the monthly closing price of major U.S. indices: Insights from artificial intelligence, time-series analysis and hybrid models. *Appl. Soft Comput.* **2018**, *71*, 685–697. [CrossRef]
19. Fama, E.F. Random walks in stock market prices. *Financ. Anal. J.* **1965**, *21*, 55–59. [CrossRef]
20. Malkiel, B.G.; Fama, E.F. Efficient capital markets: A review of theory and 810 empirical work. *J. Financ.* **1970**, *25*, 383–417. [CrossRef]
21. Kahneman, D.; Tversky, A. Prospect Theory: An Analysis of Decision under Risk. *Econometrica* **1979**, *47*, 263–291. [CrossRef]
22. Glogowski, J. Keeping Up with the Quants: Your Guide to Understanding and Using Analytics by Thomas, H. Davenport and Jinho Kim. *J. Bus. Financ. Librariansh.* **2014**, *19*, 86–89. [CrossRef]
23. Rosenberg, B.; Reid, K.; Lanstein, R. Persuasive Evidence of Market Inefficiency. *J. Portf. Manag.* **1985**, *13*, 9–17. [CrossRef]
24. Sanjoy, B. Investment Performance of Common Stocks in Relation to Their Price-Earnings Ratios: A test of the Efficient Markets Hypothesis. *J. Financ.* **1977**, *32*, 663–682.
25. Nti, I.K.; Adekoya, A.F.; Weyori, B.A. A systematic review of fundamental and technical analysis of stock market predictions. *Artif. Intell. Rev.* **2020**, *53*, 3007–3057. [CrossRef]

26. Shiller, R.J. From efficient markets theory to behavioral finance. *J. Econ. Perspect.* **2003**, *17*, 83–104. [CrossRef]
27. Thaler, R.H. The end of behavioral finance. *Financ. Anal. J.* **1999**, *55*, 12–17. [CrossRef]
28. Hsu, P.H.; Kuan, C.M. Reexamining the profitability of technical analysis with data snooping checks. *J. Financ. Econom.* **2005**, *3*, 606–628. [CrossRef]
29. Nguyen, T.H.; Shirai, K.; Velcin, J. Sentiment analysis on social media for stock movement prediction. *Expert Syst. Appl.* **2015**, *42*, 9603–9611. [CrossRef]
30. Dreman, D.N.; Berry, M.A. Overreaction, Underreaction, and the Low-P/E Effect. *Financ. Anal. J.* **1995**, *51*, 21–30. [CrossRef]
31. Anbalagan, T.; Maheswari, S.U. Classifcation and prediction of stock market index based on fuzzy metagraph. *Procedia Comput. Sci.* **2014**, *47*, 214–221. [CrossRef]
32. Ghaznavi, A.; Aliyari, M.; Mohammadi, M.R. Predicting stock price changes of tehran artmis company using radial basis function neural networks. *Int. Res. J. Appl. Basic Sci.* **2014**, *10*, 972–978.
33. Agarwal, P.; Bajpai, S.; Pathak, A.; Angira, R. Stock market price trend forecasting using machine learning. *Int. J. Res. Appl Sci. Eng. Technol.* **2017**, *5*, 1673–1676.
34. Schmeling, M. Investor sentiment and stock returns: Some international evidence. *J. Empir. Financ.* **2009**, *16*, 394–408. [CrossRef]
35. Hu, Y.; Liu, K.; Zhang, X.; Su, L.; Ngai EW, T.; Liu, M. Application of evolutionary computation for rule discovery in stock algorithmic trading: A literature review. *Appl. Soft Comput.* **2015**, *36*, 534–551. [CrossRef]
36. Khan, H.Z.; Alin, S.T.; Hussain, A. Price prediction of share market using artifcial neural network "ANN". *Int. J. Comput. Appl.* **2011**, *22*, 42–47.
37. Ballings, M.; Dirk, V.P.; Nathalie, H.; Ruben, G. Evaluating multiple classifiers for stock price direction prediction. *Expert Syst. Appl.* **2015**, *42*, 7046–7056. [CrossRef]
38. Schapire, R.E. Explaining AdaBoost. In *Empirical Inference*; Schölkopf, B., Luo, Z., Vovk, V., Eds.; Springer: Berlin/Heidelberg, Germany, 2019.
39. Su, C.H.; Cheng, C.H. A hybrid fuzzy time series model based on ANFIS and integrated nonlinear feature selection method for forecasting stock. *Neurocomputing* **2016**, *205*, 264–273. [CrossRef]
40. Padhi, D.K.; Padhy, N. Prognosticate of the financial market utilizing ensemble-based conglomerate model with technical indicators. *Evol. Intel.* **2021**, *14*, 1035–1051. [CrossRef]
41. Marwala, T.; Hurwitz, E. *Artificial Intelligence and Economic Theory: Skynet in the Market*; Springer: London, UK, 2017.
42. Malkiel, B.G. The efficient market hypothesis and its critics. *J. Econ. Perspect.* **2003**, *17*, 59–82. [CrossRef]
43. Smith, V.L. Constructivist and ecological rationality in economics. *Am. Econ. Rev.* **2003**, *93*, 465–508. [CrossRef]
44. Nofsinger, J.R. Social mood and financial economics. *J. Behav. Financ.* **2005**, *6*, 144–160. [CrossRef]
45. Bollen, J.; Mao, H.; Zeng, X. Twitter mood predicts the stock market. *J. Comput. Sci.* **2011**, *2*, 1–8. [CrossRef]
46. Henrique, B.M.; Sobreiro, V.A.; Kimura, H. Literature review: Machine learning techniques applied to financial market prediction. *Expert Syst. Appl.* **2019**, *124*, 226–251. [CrossRef]
47. Zhang, Y.; Wu, L. Stock market prediction of S&P 500 via combination of improved BCO approach and BP neural network. *Expert Syst. Appl.* **2009**, *36*, 8849–8854.
48. Chen, T.; Guestrin, C. XGBoost: A scalable tree boosting system. In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Francisco, CA, USA, 13–17 August 2016.
49. Chong, E.; Han, C.; Park, F.C. Deep learning networks for stock market analysis and prediction: Methodology, data representations, and case studies. *Expert Syst. Appl.* **2017**, *83*, 187–205. [CrossRef]
50. Avery, C.N.; Chevalier, J.A.; Zeckhauser, R.J. The CAPS prediction system and stock market returns. *Rev. Financ.* **2016**, *20*, 1363–1381. [CrossRef]
51. Prasad, V.V.; Gumparthi, S.; Venkataramana, L.Y.; Srinethe, S.; SruthiSree, R.M.; Nishanthi, K. Prediction of Stock Prices Using Statistical and Machine Learning Models: A Comparative Analysis. *Comput. J.* **2021**, 8. [CrossRef]
52. Zong, E. Forecasting daily stock market return using dimensionality reduction. *Expert Syst. Appl.* **2017**, *67*, 126–139. [CrossRef]
53. Hiransha, M.; Gopalakrishnan, E.A.; Menon, V.K.; Soman, K.P. NSE stock market prediction using deep-learning models. *Procedia Comput. Sci.* **2018**, *132*, 1351–1362.
54. Shynkevich, Y.; McGinnity, T.M.; Coleman, S.A.; Belatreche, A. Forecasting movements of health-care stock prices based on different categories of news articles using multiple kernel learning. *Decis. Support Syst.* **2016**, *85*, 74–83. [CrossRef]
55. Huang, B.; Huan, Y.; Da Xu, L.; Zheng, L.; Zou, Z. Automated trading systems statistical and machine learning methods and hardware implementation: A survey. *Enterp. Inf. Syst.* **2018**, *13*, 132–144. [CrossRef]
56. Usmani, M.; Adil, S.H.; Raza, K.; Ali, S.S.A. Stock market prediction using machine learning techniques. In Proceedings of the 2016 3rd International Conference on Computer and Information Sciences (ICCOINS), Kuala Lumpur, Malaysia, 15–17 August 2016; pp. 322–327.
57. Atsalakis, G.S.; Valavanis, K.P. Forecasting stock market short-term trends using a neuro-fuzzy based methodology. *Expert Syst. Appl.* **2009**, *36*, 10696–10707. [CrossRef]
58. Chourmouziadis, K.; Chatzoglou, P.D. An intelligent short term stock trading fuzzy system for assisting investors in portfolio management. *Expert Syst. Appl.* **2016**, *43*, 298–311. [CrossRef]
59. Arévalo, A.; Niño, J.; Hernández, G.; Sandoval, J. High-Frequency Trading Strategy Based on Deep Neural Networks. In *Intelligent Computing Methodologies*; Huang, D.S., Han, K., Hussain, A., Eds.; Springer: Berlin/Heidelberg, Germany, 2016; p. v9773.

60. Guresen, E.; Kayakutlu, G.; Daim, T.U. Using artificial neural network models in stock market index prediction. *Expert Syst. Appl.* **2011**, *38*, 10389–10397. [CrossRef]
61. Paik, P.; Kumari, B. Stock market prediction using ANN, SVM, ELM: A review. *IJETTCS* **2017**, *6*, 88–94.
62. Murekachiro, D. A review of artifcial neural networks application to stock market predictions. *Netw. Complex Syst.* **2016**, *6*, 3002.
63. Ampomah, E.K.; Qin, Z.; Nyam, G. Evaluation of Tree-Based Ensemble Machine Learning Models in Predicting Stock Price Direction of Movement. *Information* **2020**, *11*, 332. [CrossRef]
64. Leung, M.T.; Daouk, H.; Chen, A.-S. Forecasting stock indices: A comparison of classification and level estimation models. *Int. J. Forecast.* **2000**, *16*, 173–190. [CrossRef]
65. Omer, B.S.; Murat, O.; Erdogan, D. A Deep Neural-Network Based Stock Trading System Based on Evolutionary Optimized Technical Analysis Parameters. *Procedia Comput. Sci.* **2017**, *114*, 473–480.
66. Jiayu, Q.; Bin, W.; Changjun, Z. Forecasting stock prices with long-short term memory neural network based on attention mechanism. *PLoS ONE* **2020**, *15*, e0227222.
67. Xiao, Y.; Xiao, J.; Lu, F.; Wang, S. Ensemble ANNs-PSO-GA Approach for Day-ahead Stock E-exchange Prices Forecasting. *Int. J. Comput. Intell. Syst.* **2014**, *7*, 272–290. [CrossRef]
68. Huang, C.L.; Tsai, C.Y. A hybrid SOFM-SVR with a filter-based feature selection for stock market forecasting. *Expert Syst Appl.* **2009**, *36*, 1529–1539. [CrossRef]
69. Ecer, F.; Ardabili, S.; Band, S.; Mosavi, A. Training Multilayer Perceptron with Genetic Algorithms and Particle Swarm Optimization for Modeling Stock Price Index Prediction. *Entropy* **2020**, *22*, 1239. [CrossRef] [PubMed]
70. Shah, D.; Isah, H.; Zulkernine, F. Stock Market Analysis: A Review and Taxonomy of Prediction Techniques. *Int. J. Financ. Stud.* **2019**, *7*, 26. [CrossRef]
71. Zhang, G.; Patuwo, B.E.; Hu, M.Y. Forecasting with artificial neural networks: The state of the art. *Int. J. Forecast.* **1998**, *14*, 35–62. [CrossRef]
72. Basa, S.; Kar, S.; Saha, S.; Khaidem, L.; De, S. Predicting the direction of stock market prices using tree-based classifiers. *N. Am. J. Econ.* **2019**, *47*, 552–567.
73. Milosevic, N. Equity Forecast: Predicting Long Term Stock Price Movement Using Machine Learning. *arXiv* **2016**, arXiv:1603:00751.
74. Choudhury, S.S.; Sen, M. Trading in Indian stock market using ANN: A decision review. *Adv. Model. Anal. A* **2017**, *54*, 252–262.
75. Boonpeng, S.; Jeatrakul, P. Decision support system for investing in stock market by using OAA-neural network. In Proceedings of the Eighth International Conference on Advanced Computational Intelligence (ICACI), Chiang Mai, Thailand, 14–16 February 2016; pp. 1–6.
76. Yang, R.; Yu, L.; Zhao, Y.; Yu, H.; Xu, G.; Wu, Y.; Liu, Z. Big data analytics for financial Market volatility forecast based on support vector machine. *Int. J. Inf. Manag.* **2019**, *50*, 452–462. [CrossRef]
77. Yun, K.K.; Yoon, S.W.; Won, D. Prediction of stock price direction using a hybrid GA-XG Boost algorithm with a three-stage feature engineering process. *Expert Syst. Appl.* **2021**, *186*, 115716. [CrossRef]
78. Yang, B.; Zi-Jia, G.; Wenqi, Y. Stock Market Index Prediction Using Deep Neural Network Ensemble. In Proceedings of the 36th Chinese Control Conference (CCC), Dalian, China, 26–28 July 2017; pp. 3882–3887.
79. Wang, J.J.; Wang, J.Z.; Zhang, Z.G.; Guo, S.P. Stock Index Forecasting Based on a Hybrid Model. *Omega* **2012**, *40*, 758–766. [CrossRef]
80. Chenglin, X.; Weili, X.; Jijiao, J. Stock price forecast based on combined model of ARI-MA-LS-SVM. *Neural Comput. Applic.* **2020**, *32*, 5379–5388.
81. Tiwari, S.; Rekha, P.; Vineet, R. Predicting Future Trends in Stock Market by Decision Tree Rough-Set Based Hybrid System with Hhmm. *Int. J. Electron.* **2010**, *1*, 1578–1587.
82. Stankovic, J.; Markovic, I.; Stojanovic, M. Investment strategy optimization using technical analysis and predictive modeling in emerging markets. *Proc. Econ. Financ.* **2015**, *19*, 51–62. [CrossRef]
83. Atsalakis, G.S.; Valavanis, K.P. Surveying stock market forecasting techniques—Part II: Soft computing methods. *Expert Syst. Appl.* **2009**, *36*, 5932–5941. [CrossRef]
84. Bodas-Sagi, D.J.; Fernández-Blanco, P.; Hidalgo, J.I.; Soltero, D. A parallel evolutionary algorithm for technical market indicators optimization. *Nat. Comput.* **2012**, *12*, 195–207. [CrossRef]
85. de Oliviera, F.A.; Zarate, L.E.; de Azevedo Reis, M.; Nobre, C.N. The use of artificial neural networks in the analysis and prediction of stock prices. In Proceedings of the IEEE International Conference on Systems, Man, and Cybernetics, Anchorage, AK, USA, 9–12 October 2011; pp. 2151–2155.
86. Nguyen, T.-T.; Yoon, S.A. Novel Approach to Short-Term Stock Price Movement Prediction using Transfer Learning. *Appl. Sci.* **2019**, *9*, 4745. [CrossRef]
87. Bisoi, R.; Dash, P.K.; Parida, A.K. Hybrid variational mode decomposition and evolutionary robust kernel extreme learning machine for stock price and movement prediction on daily basis. *Appl. Soft Comput.* **2018**, *74*, 652–678. [CrossRef]
88. Naik, N.; Mohan, B.R. Intraday Stock Prediction Based on Deep Neural Network. *Proc. Natl. Acad. Sci. USA* **2020**, *43*, 241–246. [CrossRef]
89. Salim, L. Intraday stock price forecasting based on variational mode decomposition. *J. Comput. Sci.* **2016**, *12*, 23–27.

90. Royo, R.C.; Francisco, K. Stock market trading rule based on pattern recognition and technical analysis: Forecasting the DJIA index with intraday data. *Expert Syst. Appl.* **2015**, *42*, 5963–5975. [CrossRef]
91. Kara, Y.; Boyacioglu, M.A.; Baykan, Ö.K. Predicting direction of stock price index movement using artificial neural networks and support vector machines: The sample of the Istanbul Stock Exchange. *Expert Syst. Appl.* **2011**, *38*, 5311–5319. [CrossRef]
92. Creighton, J.; Farhana, H.Z. Towards Building a Hybrid Model for Predicting Stock Indexes. In Proceedings of the IEEE International Conference on Big Data, Boston, MA, USA, 11–14 December 2017; pp. 4128–4133.
93. Singh, J.; Khushi, M. Feature Learning for Stock Price Prediction Shows a Significant Role of Analyst Rating. *Appl. Syst. Innov.* **2021**, *4*, 17. [CrossRef]
94. Friedman, J.H. Greedy function approximation: A gradient boosting machine. *Ann. Stat.* **2001**, *29*, 1189–1232. [CrossRef]
95. Yan, D.; Zhou, Q.; Wang, J.; Zhang, N. Bayesian regularisation neural network based on artificial intelligence optimisation. *Int. J. Prod. Res.* **2017**, *55*, 2266–2287. [CrossRef]
96. Dorogush, A.V.; Ershov, V.; Gulin, A. CatBoost: Gradient boosting with categorical features support. *arXiv* **2018**, arXiv:1810.11363.
97. Patel, J.; Shah, S.; Thakkar, P.; Kotecha, K. Predicting stock and stock price index movement using Trend Deterministic Data Preparation and machine learning techniques. *Expert Syst. Appl.* **2015**, *42*, 259–268. [CrossRef]
98. Guryanov, A. Histogram-Based Algorithm for Building Gradient Boosting Ensembles of Piecewise Linear Decision Trees. In *Analysis of Images, Social Networks and Texts AIST*; Springer: Cham, Switzerland, 2019; p. 11832.
99. Xuan, H.; Lei, W.U.; YinsongYe, A. Review on Dimensionality Reduction Techniques. *Int. J. Pattern Recognit. Artif. Intell.* **2019**, *33*, 950017.
100. Sokolova, M.; Lapalme, G. A systematic analysis of performance measures for classification tasks. *Inf. Process. Manag.* **2009**, *45*, 427–437. [CrossRef]
101. TA-LIB: Technical Analysis Library. Available online: www.ta-lib.org (accessed on 10 June 2021).
102. Cai, X.; Hu, S.; Lin, X. Feature extraction using Restricted Boltzmann Machine for stock price prediction. In Proceedings of the IEEE International Conference on Computer Science and Automation Engineering, Zhangjiajie, China, 25–27 May 2012; pp. 80–83.
103. Huang, W.; Nakamori, Y.; Wang, S.Y. Forecasting stock market movement direction with support vector machine. *Comput. Oper. Res.* **2005**, *32*, 2513–2522. [CrossRef]
104. Ticknor, J.L. A Bayesian regularized artificial neural network for stock market forecasting. *Expert Syst. Appl.* **2013**, *40*, 5501–5506. [CrossRef]
105. Knoll, J.; Stübinger, J.; Grottke, M. Exploiting social media with higher-order factorization machines: Statistical arbitrage on high-frequency data of the S&P 500. *Quant. Financ.* **2019**, *19*, 571–585.
106. Lawrence, R. Using neural networks to forecast stock market prices. *Univ. Manit.* **1998**, *333*, 2006–2013.
107. Stübinger, J.; Walter, D.; Knoll, J. Financial market predictions with Factorization Machines: Trading the opening hour based on overnight social media data. *Econ. Financ. Lett.* **2019**, *5*, 28–45. [CrossRef]

mathematics

MDPI

Article

Mathematical Modeling of Manufacturing Lines with Distribution by Process: A Markov Chain Approach

Gilberto Pérez-Lechuga [1,*], Francisco Venegas-Martínez [2] and José Francisco Martínez-Sánchez [3]

1 Instituto de Ciencias Básica e Ingeniería—AAI, Universidad Autónoma del Estado de Hidalgo, Pachuca 42184, Hidalgo, Mexico
2 Escuela Superior de Economía, Instituto Politécnico Nacional, Mexico City 11350, Mexico; fvenegas1111@yahoo.com.mx
3 Escuela Superior de Apan, Universidad Autónoma de Estado de Hidalgo, Apan 42082, Hidalgo, Mexico; marzan67@gmail.com
* Correspondence: glechuga2004@hotmail.com

Abstract: Today, there are a wide variety of ways to produce goods in a manufacturing company. Among the most common are mass or line production and process production, both of which are antagonists. In an online production system, materials move from station to station, receiving added value on a well-defined layout. In a production line by process, the materials randomly visit a set of machines strategically located in order to receive a treatment, almost always through metalwork machines, according to the final product of which they will be part. In this case, there is not a predefined layout, as the incoming materials are sectioned and each piece forms a continuous flow through different workstations to receive some process. This activity depends on the function of the product and its final destination as a component of a finished product. In this proposal, Markov chain theory is used to model a manufacturing system by process in order to obtain the expected values of the average production per machine, the total expected production in all the facilities, the leisure per machine and the total productive efficiency of the system, among other indicators. In this research, we assume the existence of historical information about the use of the equipment, its failures, the causes of failure and their repair times; in any factory, this information is available in the area of manufacturing engineering and plant engineering. From this information, statistical frequency indicators are constructed to estimate transition probabilities, from which the results presented here are derived. The proposal is complemented with a numerical example of a real case obtained from a refrigerator factory established in Mexico in order to illustrate the results derived from this research. The results obtained show their feasibility when successfully implemented in the company.

Keywords: stochastic processes; Markov chains; production lines by process; flexible manufacturing systems

1. Introduction

Manufacturing is one of the most important activities in the world economy because the production of consumer goods directly influences a country's wealth, contributing to around 70% [1]. To this day, the ways of producing consumer goods have evolved in an impressive way along with the technology. Without a doubt, 5G and 4.0 technologies applied to manufacturing systems, in conjunction with hybrid optimization and artificial intelligence, make this environment a highly significant medium in modern global industry [2].

A Flexible Manufacturing System (FMS) is viewed as the integration of assembly processes, material flow, computer communications and control processes [3]. This type of structure represents the central axis around which the activities of a production line revolve. Its importance lies in exchanging machinery and processes immediately in order to manufacture a large number of different products. Usually, the topology (layout) of the plant distribution is defined through the nature of the product to be manufactured, the used materials, the machining operations, the physical arrangement of the equipment, the location of the raw material and the finished product warehouses, among other issues. The literature in this regard is quite abundant in information, although most of the authors agree on some general aspects in the ways of producing.

More specifically, the following ways of producing are common in today's manufacturing engineering:

- Unit production: It is characterized by manufacturing a single product that is made at a predetermined time according to its demand. Usually, they are unique products tailored to the client. From the perspective of large-scale manufacturing, these types of production models are of little interest.
- Batch production: It is described by the means of producing a predetermined quantity of items in a single production run. The quantity manufactured and the run times are programmed in a Master Production Schedule (MPS) according to the demand of the product. Usually, the quantity to be produced is forecast considering the available inventory using a policy (S, s). This production system becomes interesting when several batches of varied products must be manufactured. This raises the question in what order should the products be produced? From a quantitative point of view, the modeling, simulation and/or optimization of this type of system are more interesting since it involves computational problems of the NP-hard type, especially when calculating the order of production, which translates into the classic problem of sequencing operations. This scheme is especially useful in the production of seasonal goods.
- Mass production: In this system the manufacture of the product is performed under a strict order in line. At the beginning of the system is the raw material that feeds the first workstation. Product flows downstream through ρ specialized workstations increasing their added value with each visit. At the end of the line, there is a finished product warehouse where the production generated during a period of time is kept. Each workstation contributes with an incremental improvement in the functionality of the product until the final result constitutes a good that goes directly to the consumer or is a sub-assembly that will later be incorporated into another line to form part of the finished product, which can also be considered a spare part. These types of designs usually include a set of ρ (one for each machine) intermediate buffers between each workstation in order to temporarily decouple the operation of the system and thereby avoid bottlenecks or leisure times, see [3,4]. These systems are highly efficient as they generate large amounts of fully standardized product. The control of the operation can be conducted under the classic PUSH type production scheme (under a demand forecast concept) or the PULL type system using the KANBAN philosophy [5].
- Continuous production: This production system generates one or more products that generally cannot be measured in discrete units. Examples of this are the production of gasoline, milk, gas, liquors, etc. Here, production never stops and only stops when corrective and/or preventive maintenance is required on the equipment. An important characteristic of this type of production is that most of the work is performed automatically by industrial equipment with almost no human intervention.
- Production by process: This design, which is of interest in this document, is characterized by manufacturing a variety of products in quantities that can be large depending on the demand for the multiple goods to be produced. The design topology originates when there is a series of machining equipment installed in the shop in various positions (not necessarily ordered). These distributions are achieved when the machines to be used are highly automated, heavy or highly specialized, which prevents their movement (for example, electric furnaces, punching machines, cutters, shears, paint booths, etc.) and makes their movement prohibitive. Here, the product must visit the machines that will carry out the operations on the products, not necessarily in an ordered sequence.

The difficulty in mathematically modeling the latter type of process is the inherent randomness in the trajectories that the component materials (or by-products) follow. After leaving the raw material warehouse, the sub-products can follow any path required by manufacturing engineering, except the finished product warehouse. Such trajectory is determined by the design conditions required based on the model of the product to be manufactured, the quantity requested, the actual production capacity, the machining operations required, the availability of equipment and machinery, as well as the availability of tooling and labor. The uncertainty of the path that each component will follow depends on the previous state and, therefore, in this proposal such a sequence is modeled through a matrix of transition probabilities of visiting state j when it comes from state i. Hence, this model can be represented by means of a Markov chain.

Figure 1 shows two typical trajectories that could be followed by two different products within a manufacturing by process system formed by 12 workstations (placed without a specific order), a raw material warehouse and a temporary buffer or warehouse of finished products in a system in mass. That is, the finished product of this process is part of the raw material that enters another production system that is after the current process.

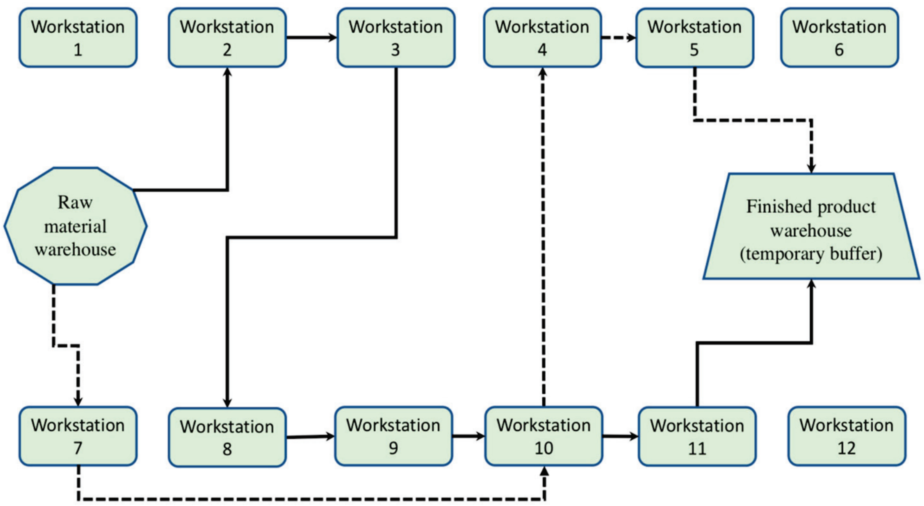

Figure 1. Two typical trajectories in a process manufacturing system. Source: Authors' own elaboration.

The objective of this research is to model this type of manufacturing process through a Markov chain in order to calculate the main efficiency and productivity indicators required by manufacturing engineering. This work extends in several aspects [6] and is distinguished with respect to the specialized literature in the following: (1) it uses Markov chain theory along with its properties and results; (2) it obtains efficiency indicators based on equipment availability; and (3) it applies the proposal to a refrigerator factory in Mexico.

The rest of the document is organized as follows. Section 2 provides a short literature review in order to establish a reference framework. Section 3 models a production line by process with a Markov chain. Section 4 illustrates the results obtained through an application developed with real information from a refrigerator manufacturing system in Mexico. Section 5 discusses the empirical results obtained. Finally, Section 6 provides the conclusions.

2. A Short Literature Review

The literature is abundant on the analysis and mathematical modeling of mass production systems since the assumptions used to build such models are easy to verify in real cases. The most representative literature of the topic of interest is mentioned below.

There are classical books that address the problem fundamentally from a stochastic perspective. Among the most popular are [7–11].

Regarding the publication of papers related to the subject, it is worth mentioning, for example, [12], which develops a model based on multi-server tandem queues with buffers and blocking, after service is developed. The model focuses on performance characteristics such as throughput and mean sojourn times. The main idea developed is the decomposition of the system into two-station subsystems using a spectral expansion method.

Similarly, Ref. [13] proposes a set of statistical indicators to evaluate the efficiency of the production line as well as the expected value of the manufactured products per unit of time, using the reliability of the installed equipment as the main element. On the other hand, Ref. [13] presents interesting results related to the pharmaceutical industry referencing the system over time. The resulting model has NP-hard characteristics, and the authors decompose the problem into two sub-problems that are easier to solve using discrete and continuous representations.

Moreover, Ref. [14] establishes a mathematical programming model that describes the behavior of a production line in a discrete way. The model specifies a multi-stage system with intermediate buffers and stochastic production times. Similarly, Ref. [15] develops a model based on the administration of tasks in hierarchical active systems including incentives.

In relation to representative models of continuous flow systems, Ref. [16] develops a mathematical programming model of discrete events in continuous time with a discrete and continuous mixed state space. The research focuses on optimally controlling the flow of the product through the system with high efficiency.

Also, an integrative approach to mathematical modeling is found in [17]. Here, the authors develop optimal manufacturing control policies for tandem production lines including KANBAN and CONWIP.

Alternative models of mathematical programming applied to the production capacity of a manufacturing line, which in their design uses optimal operating parameters of production lines, can be found in [18]. The authors in [18], develop a model to derive a production line capacity assignment problem considering an equilibrium requirement from the production in a cold rolling area. They also provide a structure in which there are several types of materials that are delivered with variable times in order to maximize the efficiency of the installed capacity in the plant.

Another interesting point of view is the optimization of the efficiency of production lines. This approach attempts to conveniently group work centers that have sequential activities into unit centers. A model associated with this approach is found in [19]. Here, a cost-oriented objective function is developed for a multiple-manned assembly line balancing problem using mixed integer mathematical programming as a tool to optimize the balancing of the production line. Finally, other applications associated with the topic can be found in [20–23].

3. The Mathematical Proposal Using Markov Chains

The use of Markov chains in various fields is very widespread. In this document, we make use of the properties of this tool to model and implement a manufacturing system by process.

3.1. Object of the Investigation

For the development of this research, the following steps are carried out:

1. Definition of the problem: The problem proposed here is a real case presented by a specialized engineering firm. The request comes from a metalworking company that manufactures different types of refrigerator models for various products. In particular, the company wants to calculate its promises to the customer based on the installed capacity in the manufacturing area in order to obtain a master production plan and calculate its plans for materials and manufacturing requirements.

2. Solution approach: Due to the complexity of calculating a quantitative model to estimate the production (avoiding the digital simulation of discrete systems) of the company in a closed way, it was determined that a Markov chain should be used to represent the dynamics of the production system by virtue of having historical information reliable to obtain relevant indicators.
3. Mathematical properties of the proposal: The definition of the matrix of transition probabilities associated with the problem uniquely characterizes the shape of the problem as well as its properties (for example, the division of communicating classes). With this, it is possible to mathematically characterize the structure by process to be used in this proposal.
4. Feasibility and relevance of the model: The feasibility and relevance of the chosen model are demonstrated in its application to the case raised by a real firm. For this reason, the set of equations that will govern the model is formally characterized through the use of statistical estimators obtained in situ.

3.2. Formalization of the Model through a Markov Chain

Let S be a finite state space. The elements of S could be vectors. Each $i \in S$ is called a state. Hereafter, $\pi = \{\pi_i \mid i \in S\}$ is a probability distribution defined in some probability space $(\Omega,\ \mathcal{F},\ \mathbb{P})$. Thus, $0\ \leq \pi_i\ \leq\ 1$ for all $i\ \in S,$ and the total mass satisfies $\sum_{i\in S}\ \pi_i = 1$. Let $X(t)$ be a random variable (vector) for each t. From now on t will denote time and it can be discrete or continuous.

The proposal developed here attempts to model a manufacturing line by process, through a stochastic process in which the probability that $X(t+\delta t)$ (position of a product at time $t+\delta t$) depends only on the previous value of $X(t)$. That is, $X(t)$ is a Markov process. Let $X(s),\ s\ \leq t$, where $X(s)$ is the history of the values of X before time t and z is the possible value of $X(t+\delta t)$. Then, it is satisfied that

$$\begin{aligned}\mathbb{P}\left[X\left(t+\delta t\right) = \right. & \left. z \mid X(s) = x(s),\ s\ \leq t\right] \\ & = \mathbb{P}\left[X(t+\delta t) = z \mid X(t) = x(t)\right].\end{aligned} \tag{1}$$

In a particular case when the process is defined in discrete times, these will be numbered by 0, 1, 2, … , or in a convenient way as the case may be. In the continuous time case, the notation $0,\ \delta,\ 2\delta,\ \ldots$, will be used, or simply δt. In the same way, the transition probability from i to j in a unit of time will be denoted by

$$P_{ij} = \mathbb{P}\left[X(t+1) = j \mid X(t) = i\right]. \tag{2}$$

In the case of a Markov chain, $X(t) = i$ denotes that the random quantity $X(t)$ is in state i. Moreover, $\pi_i\,(t) = \mathbb{P}\left[X(t) = i\right]$ satisfies the normalization equation given by $\sum_i \pi_i\,(t) = 1$. It is also satisfied the Chapman-Kolmogorov equation in the sense that $P_{kj}(s,t) = \sum_i\ P_{ji}\ (s,\ u)\ P_{kj}\ (u,t)$, for any $t > u > s\ \geq 0$ where $P_{kj}(s,t) = \mathbb{P}\left[X(t) = k \mid X(s) = j\right]$. Thus, by the Law of Total Probability and from (2), it is also true that

$$\pi_i\,(t+1) = \sum\nolimits_j P_{ij}\,\pi_j\,(t) \text{ with } \sum\nolimits_i \pi_i\,(t) = 1. \tag{3}$$

with $P_{ij} = \mathbb{P}\left(X(1) = j \mid X(0) = i\right)$. Equivalently, we may write

$$\boldsymbol{\pi^t} = \begin{pmatrix} \pi_1\,(t) \\ \pi_2\,(t) \\ \vdots \\ \pi_n\,(t) \end{pmatrix},\ \boldsymbol{P} = \begin{pmatrix} P_{11} & \cdots & P_{1n} \\ \vdots & \ddots & \vdots \\ P_{n1} & \cdots & P_{nn} \end{pmatrix} \text{ and } \boldsymbol{\vartheta} = \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}$$

where

$$\boldsymbol{\pi^{t+1}} = \boldsymbol{P\pi^t}\,,\ \boldsymbol{\vartheta^T\pi^t} = 1 \tag{4}$$

and

$$\boldsymbol{\pi^t} = \boldsymbol{P^t}\boldsymbol{\pi^0},\ \boldsymbol{P\vartheta} = \boldsymbol{\vartheta}\ \text{(each rowof } \boldsymbol{P} \text{ sums 1)}. \tag{5}$$

In addition, the following conditions are satisfied in steady state

$$\pi_i = \lim_{\ t\ \to\infty}\pi_i\ (t) \tag{6}$$

if it exists, and the transition equations

$$\pi_i = \sum\nolimits_j P_{ij}\ \pi_j, \tag{7}$$

as well as the balance equations in steady state hold

$$\pi_m\ \sum\nolimits_{m,\ m\neq i} P_{mi} = \sum\nolimits_{i,\ i\neq j} P_{ij}\ \pi_i \tag{8}$$

The corresponding normalization equation is given by

$$\sum\nolimits_i \pi_i\ (t) = 1 \tag{9}$$

The steady-state balance equations intuitively mean that the average number of transitions into some state per unit of time must be equal to the average number of transitions from the same state to other states [8].

In the matrix-vector form, it is written that in the steady state $\boldsymbol{\pi}$ = lim $_{t\ \to\infty}$ $\boldsymbol{\pi^t}$ if the limit exists. Moreover, $\boldsymbol{P} = \boldsymbol{\pi}\,\boldsymbol{P}$, and the normalization equation is $\boldsymbol{\vartheta^T\pi} = 1$. Analogously, $\boldsymbol{\pi}$ = lim$_{t\ \to\infty}$ $\boldsymbol{P^t}\ \boldsymbol{\pi^0}$ provided the limit exists.

Therefore, the characteristics of a production system by process allow it to be seen as a Markov chain since the passage of the material towards the state $X(t+1)$, given that it is currently at state $X(t)$, only depends on the previous state. Here, the process is defined through the path that materials follow in the form of semi-finished products to be machined once they leave the raw material warehouse. The process ends when the finished product is in a temporary buffer (generally as a partial element of other products or as a spare part) to be assigned to a new assembly line or stored for sale.

Thus, at each step, the material is necessarily in some state $X(t)$, and will evolve into the state $X(t+1)$, according to the machining requirements demanded by it.

Once the above is formalized, the modeling of this type of production line is performed under the following considerations:

1. All the material entering the system comes from a single source called the raw material warehouse and will be denoted as the initial state s_1.
2. The set $S' \subseteq S$ defined as $S' = \{s_2,\ \ldots,\ s_{\rho-1}\}$ represents the intermediate stages of the manufacturing process. That is, where the material is machined to add value to become a by-product of the process. In principle, any $s_j \in S'$ is accessed from $s_i \in S'$.
3. For all $s_i\ \in S'$, there is a $s_j\ \in S'$ with $i \neq j$, such that $s_i\ \longleftrightarrow\ s_j$. This means that any material in the manufacturing process can return with positive probability to any of the previous states. For practical purposes this is called material reprocessing.
4. The state s_ρ is the only one that allows material to exit the system. This represents the finished product warehouse and can be accessed from any $s \in S'$. From here, the materials move onto a general assembly line in series and never return to any of the previous nodes.
5. The material from the state s_1 is sent to any $s \in Z'$. This constitutes the product or productive unit of the system on which operations will be carried out in the rest of the states $s\ \in S'$.
6. The transition matrix P will be assumed to be known. Transition probabilities are easily retrievable from the plant engineering and manufacturing department historical files.

7. The reliability of an equipment (the probability that it will function when required) located in the state $s \in Z'$ in the instant t, is given by the function

$$R(t) = \mathbb{P}(\mathcal{T} \geq t) = \int_t^\infty f_{\mathcal{T}}(t)\, dt, \quad \mathcal{T} \geq 0 \tag{10}$$

where $f_{\mathcal{T}}(t)$ is the failure density function of the equipment at the states $s \in Z'$, and $\mathcal{T}$ denotes the random variable that represents the instant where a failure occurs. Here, it will be assumed that

$$\mathrm{E}(\mathcal{T}) = \int_0^\infty t\, d\, F_{\mathcal{T}}(t) < \infty. \tag{11}$$

Hereinafter, E denotes the mathematical expectation.

8. Historical information is available on the machining equipment and the number of operations carried out on the materials that visit them.

3.3. Characterization of the Model

During our analysis, the state space $S \subset \mathbb{R}^\rho$ is given by the diverse workstations to which the product can arrive during its trajectory. The analyzed system is made up of the states $s_1, s_2, \ldots s_\rho$, where s_1 represents the raw material warehouse and the state s_ρ represents the finished product warehouse. So, a typical transition matrix would be given by Table 1.

Table 1. Typical matrix of transition probabilities used in this process. Source: Authors' own elaboration.

	s_1	s_2	$\cdots$	$s_{\rho-1}$	s_ρ
s_1	0		$\cdots$	$p_{1,\rho-1}$	0
s_2	0	0	$\cdots$	$p_{2,\rho-1}$	$p_{2,\rho}$
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$
$s_{\rho-1}$	0	$p_{\rho-1,2}$	$\cdots$	0	$p_{\rho-1,\rho}$
s_ρ		0	$\cdots$		1

According to Table 1, $P_{ij} \geq 0$ and $\sum_{j=1}^{\rho} P_{ij} = 1, \forall\, i$. This research uses the notation $s_i \to s_j$ to express that a state s_j is accessible from a state s_i. In our case, the initial state s_1 represents the raw material warehouse and, therefore, it is a non-return state. In turn, the final state s_ρ represents the finished product warehouse and once the product reaches it, it remains there until it is required later on in another production line. s_ρ is a buffer between the first manufacturing process and the main assembly line. This is an absorbing and therefore recurrent state. The intermediate states $s_2, \ldots, s_{\rho-1}$, represent the existing machining teams between the raw material warehouse and the finished product warehouse. From the characteristics of the transition matrix in Table 1, the following results from [23–25] hold.

Let $i = 1, \ldots, \rho$ be the set of states in which a by-product can be found within the production line. If the process $X(\cdot)$ begins with a known initial distribution $\pi = (\pi_1, \ldots, \pi_\rho)$, i.e.,

$$\mathbb{P}(X(1) = s_j) = \sum_{i=1}^{\rho} \mathbb{P}(X(1) = s_j \mid X(0) = s_i)\, \mathbb{P}(X(0) = s_i) = \sum_{i=1}^{s} \pi_i P_{ij},$$

then the distribution of $X(1)$ is defined through $\boldsymbol{\pi P}$ and, in general, the distribution of $X(t)$ is defined through $\boldsymbol{\pi P^t}$. That is,

$$\lim_{t\to\infty} \boldsymbol{\pi}\, \boldsymbol{P^t} \Rightarrow \boldsymbol{\pi^t}$$

where $\boldsymbol{P^0} = \boldsymbol{I}$ is the identity matrix. Here, the symbol $\Rightarrow$ means weak convergence and equality $\boldsymbol{\pi^t} = \boldsymbol{\pi}\,\boldsymbol{P^t}$ is equivalent to $\boldsymbol{\pi^t} = \boldsymbol{P^T}\,\boldsymbol{\pi^{t-1}}$, where $\boldsymbol{P^T}$ means $\boldsymbol{P}$ transposed.

During the manufacturing process, each state commutes with the rest. That is, the state s_j is reachable from the state s_i and vice versa (denoted as $s_i \longleftrightarrow s_j$). Therefore, this relation dissects the state space into a family of classes G that commute mutually. In practice, a commutative relationship between two states, s_i and s_j, means that a material in the state s_j can go back to be reprocessed to the previous station s_i except with the raw material warehouse s_1 with positive probability. Similarly, a material in the state s_ρ cannot return to any other state $s_1, s_2, \ldots, s_{\rho-1}$.

In decomposing the Markov chain into classes G_{s_i}, with $s_i \in S$, the following classes are clearly distinguished:

1. G_{s_1} is a class with only one non-return state s_1 or transcendent state since $\mathbb{P}\left(X(t) = s_\rho\right) = 0$, for $t > 1$ (assumption 1).
2. G_{s_j} is a recurrent communicating class since $s_i \longleftrightarrow s_j, \forall\, s_i, s_j \in S'$. This means that $\mathbb{P}\left(X(t) = s_j\right) = 1,$ for $t > 1$ (assumption 2).
3. G_{s_ρ} is a communicating class with a single absorbing state (assumption 3).

As G_{s_ρ} is a closed communicating class with a single state s_ρ, then for $P^t_{\rho\rho} = 1$, for every $t = 1, 2, \ldots$. Hence s_ρ is an absorbing state and, therefore, it is recurrent.

Classes satisfy the relation $G_{s_1} \prec G_{s_j} \prec G_{s_\rho}$. That is, given $s_1 \in G_{s_1}$ and $s_j \in G_{s_j}$, $s_1 \rightarrow s_j$, also $s_j \rightarrow s_\rho$. Therefore, G_{s_1} and G_{s_j} constitute a transcendent class while G_{s_ρ} is a maximal class since the state s_ρ is absorbent.

Due to the assumption that a material processed in the state $s \in S'$ can return to be reprocessed on previous workstations or be transferred to higher states, it is satisfied that for all $s_i \in \{ S' \cup s_\rho \}$.

$$\mathbb{P}\left[X(t+\delta t) = \{s_i \cup s_\rho\} \,\middle|\, X(t) = s_1\right] = 1, \tag{12}$$

Thus, the states $s \in S'$ are recurrent and, therefore, the class G_{s_j} is recurrent. Consequently, for all $s_i \in S'$, and s_ρ, there is a matrix v_{ij} such that

$$v_{ij} = \sum_{t=0}^{\infty} P\left[X(t+\delta t) = s_j \,\middle|\, X(t) = s_i\right] = \sum_{t=0}^{\infty} P^t_{ij}. \tag{13}$$

Equation (13) defines the expected number of visits of the process to state s_j when it starts in s_i. In practical terms, this represents the expected number of demands per unit of time that each workstation on the manufacturing line has during the planning horizon.

The duration of a visit in each state i in the instant t is defined by the random variable L^t_i and this is formally assigned an exponential distribution. Its meaning can be explained as the time required by each workstation to complete an operation on the visiting product. In order to calculate the duration of the visit to each state in the chain, it is defined

$$\tau = \inf\ \{t \geq 1 \,|\, X(t) = s\} \tag{14}$$

where $\tau^0_i = 0$, $\tau^1_i = \tau_i$, and, in general, $\tau^{t+1}_i = \inf\left\{t \geq \tau^t_i + 1 \,\middle|\, X(t) = i\right\}$. Then, the duration of the t-th visit is given by

$$L^t_i = \begin{cases} \tau^t_i - \tau^{t-1}_i, & if\ \tau^{t-1}_i < \infty, \\ 0, & \text{otherwise.} \end{cases}$$

In this proposal, the random variable $L^t_i \sim (\tau^{-1})$ where for a set of γ samples an unbiased estimator for τ is given by

$$\hat{\tau} = \sum_{t=1}^{\gamma} \left(\tau^t_i - \tau^{t-1}_i\right) \gamma^{-1}.$$

The following result is relevant for the analysis of this proposal [14]. For $t = 2, 3, \ldots,$ conditional on $\tau_i^{t-1} < \infty,\ L_i^t$ is independent of $\left\{ X(t) \,\middle|\, t \leq L_i^{t-1} \right\}$ and

$$\mathbb{P}\left(L_i^t = k \,\middle|\, \tau_i^{n-1} < \infty\right) = P\left[\,X(t)\right] = k. \tag{15}$$

Regarding the conditional probabilities of ever visiting the state k, given that the Markov chain was initially at state j, and the conditional probability of an infinite number of visits to the state k given that the Markov chain was originally in the state j will be denoted according to conventional literature as f_{jk} and g_{jk}. The value of f_{jk} is important as it represents the conditional probability of ever visiting the state k, given that the Markov chain was initially at state j. In this proposal this indicator will be used as an estimator of the productivity of the corresponding workstation.

Formally, the conditional probability f_{jk} that the first passage from j to k occurs in exactly t steps is given by

$$f_{jk} = \mathbb{P}\left[N_k(\infty) - N_k\,(t) > 0 \,|\, X(t) = j\right] = \sum_{t=1}^{\infty} f_{jk}\,(t) \tag{16}$$

where the random variable $N_k(\infty) = \lim_{t\to\infty} N_k\,(t)$ represents the total occupation time of the state k. The conditional probability of infinitely many visits to the state s, given that the Markov chain was at state j initially, satisfies

$$g_{jk} = \mathbb{P}\left[N_k\,(\infty) - N_k\,(t) = \infty \,|\, X\,(t) = j\right] = f_{jk}\, g_{kk} \tag{17}$$

where

$$g_{kk} = \lim_{t\to\infty} (f_{kk})^t \tag{18}$$

and

$$N_k(\infty) = \textstyle\sum_{m=1}^{\infty} Z_k\,(m),\ N_k(n) = \sum_{m=1}^{n} Z_k\,(m), \tag{19}$$

where

$$Z_k\,(n) = \begin{cases} 1, \text{ if } X(t) = k, \\ 0, \text{ if } X(t) \neq k. \end{cases} \tag{20}$$

Similarly, we denote by $f_{jk}\,(t)$ the conditional probability that the first passage from j to k occurs in exactly t steps, i.e.,

$$f_{jk}\,(t) = \mathbb{P}\left[\,V_k(t) \,|\, X(0) = j)\right]. \tag{21}$$

This magnitude makes sense in this analysis as it represents the time spent in a by-product visiting a workstation for the first time

$$V_k(t) = [X(t) = k,\ \ X(m) \neq k] \text{ for } \ m = 1, 2, \ldots, t-1. \tag{22}$$

As a consequence, from (21), the probability of ever visiting state k when coming from state j is given by

$$f_{jk} = \sum_{t=1}^{\infty} f_{jk}\,(t). \tag{23}$$

According to (22), $V_k(t)$ is the event of visiting for the first time among the times 1, 2 . . . , at which state k is visited at time t, see [24]. Two simple extensions of Equation (21) are the probability of never visiting the state k when the process is in j, $1 - f_{jk}$, as well as the probability of never returning to the state k after leaving it, $1 - f_{kk}$.

Notice now that, from Equations (17) and (18), it can be readily seen

$$g_{kk} = \begin{cases} 1, \text{ iff } f_{kk} = 1 \\ 0, \text{ iff } f_{kk} < 1 \end{cases} \tag{24}$$

and

$$f_{kk} = \begin{cases} < 1, & \text{iff } \sum\limits_{t=1}^{\infty} p_{kk}(t) < \infty, \\ = 1, & \text{iff } \sum\limits_{t=1}^{\infty} p_{kk}(t) = \infty. \end{cases} \tag{25}$$

In particular, for the matrix shown in Table 1, $f_{jk} = 1$ for $j, k \in Z'$ if $j \longleftrightarrow k$. Otherwise, $f_{jk} = 0$. Similarly, for $j \in Z'$ and $k = s_1$, $f_{jk} = 0$.

From the above, given Z and the absorbing state s, and for recurrent state j, the absorption probabilities satisfy the following system of equations.

$$f_{js} = \sum_{i \in Z} p_{ji} f_{is},\ j \in Z f_{ss} = 1, \quad f_{js} = 0,\ j \neq s.$$

Let the matrix of vectors be $\boldsymbol{Q}_\alpha$ defined as

$$\boldsymbol{Q}_\alpha = \left(\boldsymbol{P}^1, \boldsymbol{P}^2, \ldots, \boldsymbol{P}^\alpha, \boldsymbol{P}^\rho\right), \tag{26}$$

where $\boldsymbol{P}^\alpha = \boldsymbol{\theta} \in \mathbb{R}^\rho$, that is, the matrix $\boldsymbol{Q}_\alpha$ results from substituting the α-th position of the matrix $\boldsymbol{P}$ by the zero vector $\boldsymbol{\theta}$. Then, by Equation (16), it is satisfied

$$f_{jk}(t) = \boldsymbol{Q}_\alpha f_{jk}(t-1),\ t = 2, 3, \ldots \tag{27}$$

Equation (27) represents the probability of visiting the state k in t steps given that, initially, the process was at the state j.

Finally, an important equation for the limiting probabilities of the process is given by the expected number of steps t required to return to state k. Formally,

$$m_{kk} = E(t) = \sum_{t=1}^{\infty} t\ f_{kk}(t). \tag{28}$$

Let Ψ be the set of non-recurring states in Z, and let $Q = \left\{P_{jk} : j, k \in \Psi\right\}$ be the matrix of transition probabilities of the states that map from Ψ to Ψ. Then, the vector matrix of transition probabilities of the states that map from Ψ to Ψ. Then, the vector $\boldsymbol{m}$ of mean absorption times of the chain associated with the manufacturing process can be obtained from the equality

$$\boldsymbol{m} = (I - Q)^{-1}\mathbf{1} \tag{29}$$

where $\mathbf{1}$ is the column vector whose components are ones.

Now, we focus on the conditions of availability of the equipment in each workstation when receiving a visit from a by-product. From Equation (10), the Mean Time to Failure (MTTF) of each equipment is defined as the expectation of the random variable $\mathcal{T}$ as follows

$$\text{MTTF} = \text{E}(\mathcal{T}) = -\int_0^\infty t\frac{dR}{dt}\,dt = \int_0^\infty R(\mathcal{T})\,dt \approx \lambda^{-1}. \tag{30}$$

In the same way, let $m(t)\,\Delta t = P\,(t \leq \mathcal{T} \leq t + \Delta t)$ be the probability that the repair of failed equipment requis a time between $t + \Delta t$ to be repaired. Then, the Mean Time to Repair (MTTR) is given by

$$\text{MTTR} = \int_0^\infty t\,m(t)\,dt \ \approx\ \mu^{-1}. \tag{31}$$

Regardless of the Probability Density Function (PDF) of the random variable $\mathcal{T}$, the λ^{-1} and μ^{-1} values represent the MTTF and the MTTR, respectively. In turn, these are used as estimators of availability and unavailability (due to repairs) of equipment.

Define now the functions

$$P_{i,j} = P_{1,0} = \begin{cases} \lambda, & if\ i \neq j \\ 1-\lambda, & if\ \ i = j \end{cases} \quad (32)$$

$$R_{j,i} = R_{0,1} = \begin{cases} \mu,\ if\ \ i \neq j \\ 1-\mu,\ if\ \ \ j = j \end{cases}$$

where i and j take the value 1 if the machine located at station s is available, and 0 otherwise. Let $\pi(\alpha, t)$ be the probability of being in the α-state then, under steady state conditions it is satisfied that [3,8]

$$\pi(0,t) = \lim_{t\to\infty}\left[\pi\,(0,0)\,(1-\lambda-\mu)^t + \frac{\lambda}{\mu+\lambda}\left[1-(1-\lambda-\mu)^t\right]\right]\frac{\lambda}{\mu+\lambda} \quad (33)$$

$$\pi(1,t) = \lim_{t\to\infty}\left[\pi\,(1,0)\,(1-\lambda-\mu)^t + \frac{\mu}{\mu+\lambda}\left[1-(1-\lambda-\mu)^t\right]\right] = \frac{\mu}{\mu+\lambda} \quad (34)$$

Here, $\eta_j = \pi(1,t)_j$ represents the efficiency of equipment located at state s_j. Similarly, if the nominal capacity of that equipment is $\mathcal{Q}$ pieces per unit of time, then the average production rate is given by:

$$\overline{\mathcal{Q}}_j = \mathcal{Q}\,\eta_j,\ j = 1, 2, \ldots, \rho \quad (35)$$

The expected productivity associated with the model can be estimated by (in pieces per unit of time)

$$\mathcal{P}_j^{Tot} = \sum_{i=1}^{\rho} \overline{\mathcal{Q}}_j f_{ij}\ \forall j = 1, 2, \ldots, \rho \quad (36)$$

Finally, another relevant measure of the system is given by its leisure (or period of inactivity due to equipment stoppage). Leisure represents the productivity lost by not having the equipment time to do it. This measure can be approximated by the expression

$$\mathcal{L}_j = 1 - \pi(1,t) = \frac{\lambda}{\mu+\lambda}. \quad (37)$$

Below is an application of the previous concepts.

4. An Application to a Refrigerator Factory

Metalwork machines are devices that are driven by electricity to perform different operations on metals such as sanding, knurling, drilling, boring, facing, threading and turning, among others. The most frequently used are: lathe machine, milling machine, grinding machine, drilling machines, punching machines, bending machines, paint booths, welding stations, shaper machines, broaching machines, saw machines, planer machine, shearing machine, hobbing machines, drill press and more.

To illustrate the use of the proposed technique, consider an integrated process production section of 12 metalwork machines, a raw material warehouse (steel or aluminum sheets) and a temporary buffer where finished products are placed before being used in another assembly line as in Figure 1. The matrix of transition probabilities used in this example is provided in Table 2. In this case, state 1 represents the raw material warehouse from where the material that enters the manufacturing process comes, and state 12 is the finished product warehouse where all production is accumulated before entering the assembly line. Transition probabilities were obtained from historical data generated in manufacturing engineering and plant engineering. The first defines the machining operations to be carried out on each product, and the second provides information about the availability, capacity and efficiency of the equipment.

Table 2. Transition matrix P_{ij} used in the numerical example. Source: Authors' own elaboration.

States	1	2	3	4	5	6	7	8	9	10	11	12
1	0.0000	0.0870	0.0563	0.0599	0.0011	0.0448	0.0876	0.0807	0.0000	0.0138	0.5689	0.0000
2	0.0000	0.0000	0.0708	0.0424	0.0153	0.1176	0.0464	0.0788	0.0153	0.0616	0.1101	0.4418
3	0.0000	0.0619	0.0000	0.1064	0.0000	0.0687	0.0090	0.0000	0.0166	0.0981	0.0000	0.6392
4	0.0000	0.0375	0.1004	0.0000	0.0554	0.0551	0.0764	0.1163	0.0000	0.0656	0.0806	0.4127
5	0.0000	0.0228	0.1070	0.0158	0.0000	0.1167	0.0065	0.0596	0.1088	0.0943	0.0911	0.3775
6	0.0000	0.0912	0.1235	0.0087	0.0255	0.0000	0.0455	0.0908	0.0094	0.0339	0.0773	0.4941
7	0.0000	0.0830	0.0000	0.1126	0.0293	0.0454	0.0000	0.0000	0.0953	0.0251	0.1001	0.5093
8	0.0000	0.0541	0.1195	0.1035	0.0520	0.1244	0.0458	0.0000	0.0644	0.0376	0.0167	0.3819
9	0.0000	0.0739	0.0343	0.1215	0.0466	0.0000	0.1177	0.0999	0.0000	0.0230	0.0228	0.4603
10	0.0000	0.0000	0.0000	0.1130	0.0818	0.0811	0.0878	0.0000	0.0000	0.0000	0.0000	0.6363
11	0.0000	0.0262	0.0115	0.0000	0.0194	0.0724	0.0605	0.1185	0.1000	0.0978	0.0000	0.4492
12	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	1.0000

The steady state transition matrix $\boldsymbol{P^t}$obtained from the numerical example is achieved for $t = 30$, that is, $\boldsymbol{P}^{30}$.

Table 3 shows the Matrix v_{ij}, i.e., the expected number of finished products entering the final warehouse, in Equation (13). In this case, state 12 has been selected because it is the most representative of the system, since it is the connection between two different areas in the manufacturing process. In this case, manufacturing by process and mass manufacturing. Note that the last column (state 12) reflects the expected total quantity of finished goods arriving at the finished goods warehouse.

Table 3. Matrix v_{ij} , the expected number of visits to each state of the process. Source: Authors' own elaboration.

States	1	2	3	4	5	6	7	8	9	10	11	12
1	0.0000	0.1741	0.1523	0.1602	0.0639	0.1866	0.1930	0.2215	0.1112	0.1413	0.6468	27.1230
2	0.0000	0.0556	0.1354	0.1088	0.0531	0.1891	0.1013	0.1410	0.0605	0.1216	0.1583	28.6759
3	0.0000	0.0916	0.0456	0.1461	0.0295	0.1156	0.0505	0.0447	0.0348	0.1322	0.0401	29.2213
4	0.0000	0.0929	0.1659	0.0776	0.0931	0.1400	0.1287	0.1720	0.0523	0.1302	0.1333	28.6471
5	0.0000	0.0833	0.1729	0.0975	0.0429	0.1910	0.0741	0.1288	0.1487	0.1571	0.1398	28.5885
6	0.0000	0.1353	0.1761	0.0763	0.0561	0.0767	0.0911	0.1403	0.0511	0.0931	0.1221	28.8286
7	0.0000	0.1274	0.0615	0.1663	0.0647	0.1111	0.0600	0.0746	0.1321	0.0820	0.1523	28.7764
8	0.0000	0.1156	0.1931	0.1763	0.0907	0.2000	0.1060	0.0732	0.1041	0.1063	0.0815	28.6536
9	0.0000	0.1255	0.1035	0.1909	0.0853	0.0830	0.1691	0.1598	0.0488	0.0842	0.0869	28.7563
10	0.0000	0.0395	0.0526	0.1505	0.1061	0.1285	0.1211	0.0479	0.0338	0.0423	0.0498	29.1678
11	0.0000	0.0780	0.0738	0.0767	0.0596	0.1392	0.1167	0.1692	0.1371	0.1424	0.0486	27.5663
12	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	30.0000

Derived from the results of Tables 3 and 4, we show the probabilities $f_{jk}\ (t)$ for $k = 12,\ t = 1,\ 2,\ \ldots,\ 12$, i.e., the conditional probability f_{jk} that the first passage from j to k occurs in exactly t steps (as in Equation (27)).

Table 4. Probability matrix $f_{jk}\,(t)$ for $j = 1, \ldots, 12,\ k = 12,\ t = 1,\ 2,\ \ldots,\ 12$. Source: Authors' own elaboration.

States	t = 1	t = 2	t = 3	t = 4	t = 5	t = 6	t = 7	t = 8	t = 9	t = 10	t = 11	t = 12
1	0.4614	0.2611	0.1337	0.0699	0.0359	0.0185	0.0095	0.0049	0.0025	0.0013	0.0007	0.0003
2	0.2760	0.1371	0.0704	0.0363	0.0187	0.0096	0.0049	0.0025	0.0013	0.0007	0.0003	0.0002
3	0.1799	0.0872	0.0457	0.0233	0.0120	0.0062	0.0032	0.0016	0.0008	0.0004	0.0002	0.0001
4	0.2902	0.1449	0.0738	0.0381	0.0196	0.0101	0.0052	0.0027	0.0014	0.0007	0.0004	0.0002
5	0.3197	0.1455	0.0764	0.0393	0.0202	0.0104	0.0054	0.0028	0.0014	0.0007	0.0004	0.0002
6	0.2510	0.1246	0.0631	0.0327	0.0168	0.0086	0.0044	0.0023	0.0012	0.0006	0.0003	0.0002
7	0.2214	0.1304	0.0674	0.0347	0.0179	0.0092	0.0047	0.0024	0.0013	0.0006	0.0003	0.0002
8	0.3085	0.1508	0.0771	0.0396	0.0204	0.0105	0.0054	0.0028	0.0014	0.0007	0.0004	0.0002
9	0.2453	0.1434	0.0733	0.0377	0.0194	0.0100	0.0051	0.0026	0.0014	0.0007	0.0004	0.0002
10	0.1623	0.0987	0.0498	0.0256	0.0132	0.0068	0.0035	0.0018	0.0009	0.0005	0.0002	0.0001
11	0.2662	0.1361	0.0722	0.0370	0.0191	0.0098	0.0051	0.0026	0.0013	0.0007	0.0004	0.0002
12	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000

Table 5 shows the results of the probability f_{jk} of ever visiting state k when coming from state i, as in Equation (23).

Table 5. Probability matrix f_{jk} for the proposed instance. Note that $g_{kk} = 0,\ \forall\, k$. Source: Authors' own elaboration.

States	1	2	3	4	5	6	7	8	9	10	11	12
1	0.0000	0.0872	0.0895	0.0890	0.0603	0.1286	0.0946	0.1257	0.1060	0.1217	0.0479	0.9712
2	0.0000	0.0556	0.0588	0.0586	0.0357	0.0580	0.0493	0.0525	0.0424	0.0550	0.0409	0.5512
3	0.0000	0.0298	0.0436	0.0292	0.0283	0.0386	0.0387	0.0416	0.0165	0.0287	0.0382	0.3590
4	0.0000	0.0554	0.0583	0.0721	0.0339	0.0750	0.0451	0.0439	0.0498	0.0592	0.0465	0.5814
5	0.0000	0.0606	0.0585	0.0748	0.0412	0.0607	0.0634	0.0604	0.0330	0.0564	0.0422	0.6163
6	0.0000	0.0442	0.0450	0.0621	0.0284	0.0713	0.0405	0.0398	0.0394	0.0554	0.0390	0.5005
7	0.0000	0.0444	0.0589	0.0418	0.0329	0.0577	0.0567	0.0695	0.0306	0.0536	0.0451	0.4839
8	0.0000	0.0615	0.0652	0.0602	0.0350	0.0614	0.0543	0.0682	0.0348	0.0644	0.0611	0.6144
9	0.0000	0.0517	0.0647	0.0557	0.0353	0.0771	0.0419	0.0490	0.0465	0.0577	0.0601	0.5358
10	0.0000	0.0395	0.0503	0.0267	0.0199	0.0383	0.0265	0.0446	0.0322	0.0406	0.0474	0.3615
11	0.0000	0.0518	0.0591	0.0713	0.0378	0.0569	0.0497	0.0392	0.0307	0.0388	0.0463	0.5041
12	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000

Finally, vector $\boldsymbol{m}$ of mean absorption times associated with the system from state 2 to state 11 is shown in Table 6, as in Equation (29). Note that station 1 has a zero-value associated with it because, although its work is important, it does not contribute to the added value of the product since it is only a temporary warehouse for raw materials. Similarly, workstation 12 represents a point of accumulation of material of the entire manufacturing line, thus the quantities obtained represent the average values of accumulation of material per unit of time.

Table 6. Mean absorption time $(\boldsymbol{m})$, states 2 to 11. Source: Authors' own elaboration.

2	3	4	5	6	7	8	9	10	11
2.9924	2.1727	1.8129	2.2781	2.3696	2.1348	2.0858	2.3853	2.1720	1.7708

5. Discussion of Empirical Results

In this section, empirical results are presented and discussed. Table 7 shows the average production per workstation (states) pieces/unit of time. The estimators in Table 7 are part of other necessary calculations, such as material requirement planning (MRP), and manufacturing requirement planning (MRP II). Likewise, the estimators found allow us to approximate a Master Production Schedule (MPS) in order to create commitments with customers and suppliers.

Table 7. Average production per workstation (states) pieces/unit of time.

Row	Station Measure	1	2	3	4	5	6	7	8	9	10	11	12
1	μ	2.0000	3.1824	1.6939	2.1028	3.9182	1.2838	0.8273	0.7283	0.9273	1.4828	1.2303	0.9891
2	λ	0.0100	0.0300	0.2091	0.1912	0.0245	0.0439	0.0234	0.0129	0.0121	0.3726	0.0123	0.8373
3	$\mathcal{Q}$	14	16	12	15	14	17	16	14	12	16	18	14
4	η	0.9950	0.9907	0.8901	0.9166	0.9938	0.9669	0.9725	0.9826	0.9871	0.7992	0.9901	0.5416
5	$\overline{\mathcal{Q}}$	13.9303	15.8506	10.6815	13.7496	13.9130	16.4378	15.5599	13.7563	11.8454	12.7869	17.8218	7.5818
	Station	1	2	3	4	5	6	7	8	9	10	11	12
6	1	0.0000	1.3815	0.9555	1.2234	0.8393	2.1139	1.4726	1.7294	1.4892	1.5566	0.8534	7.3634
7	2	0.0000	0.8813	0.6278	0.8053	0.4972	0.9541	0.7666	0.7228	0.6224	0.7037	0.7283	4.1790
8	3	0.0000	0.4716	0.4658	0.4013	0.3937	0.6352	0.6016	0.5726	0.4931	0.3669	0.6808	2.7216
9	4	0.0000	0.8789	0.6229	0.9916	0.4721	1.2321	0.7013	0.6037	0.5198	0.7575	0.8293	4.4080
10	5	0.0000	0.9599	0.6246	1.0288	0.5739	0.9983	0.9872	0.8313	0.7158	0.7215	0.7516	4.6725
11	6	0.0000	0.7006	0.4805	0.8541	0.3945	1.1720	0.6301	0.5480	0.4719	0.7082	0.6959	3.7947
12	7	0.0000	0.7044	0.6291	0.5748	0.4572	0.9488	0.8827	0.9555	0.8228	0.6849	0.8040	3.6690
13	8	0.0000	0.9747	0.6962	0.8274	0.4875	1.0090	0.8445	0.9386	0.8083	0.8237	1.0884	4.6586
14	9	0.0000	0.8191	0.6915	0.7662	0.4910	1.2681	0.6522	0.6739	0.5803	0.7379	1.0704	4.0623
15	10	0.0000	0.6261	0.5374	0.3677	0.2774	0.6293	0.4117	0.6136	0.5284	0.5189	0.8456	2.7405
16	11	0.0000	0.8206	0.6314	0.9797	0.5262	0.9361	0.7732	0.5387	0.4638	0.4964	0.8256	3.8223
17	12	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
18	Average produtivity	0.0000	9.2186	6.9628	8.8204	5.4100	11.8970	8.7238	8.7281	7.5157	8.0762	9.1733	46.0917
19	$\mathcal{L}$	0.0000	0.0093	0.1099	0.0834	0.0062	0.0331	0.0275	0.0174	0.0129	0.2008	0.0099	0.4584

In this proposal, the probabilities of ever visiting each station have been used as an estimator of the probability of arriving at that place. The assumption of independence between the probabilities of visiting and availability of each station, as well as the previous knowledge of the nominal capacity of the equipment, allows the estimation of production by work area. An important aspect of the proposal is the shape of the transition matrix $\boldsymbol{P}$. This is common in processes of the same type; therefore, it facilitates the characterization of these systems and, therefore, their analysis from a quantitative perspective.

With this proposal, several indicators can be obtained from the basic knowledge of the $\boldsymbol{P}$ matrix and the initial distribution $\boldsymbol{\pi}$. Such statistics are normally available in the area of engineering and maintenance departments of the majority metalworking industries.

The way to obtain them and their practical meaning is as follows (The number assigned to each line corresponds to the explanation given below):

1. The 12 stations that make up this study are defined.
2. Here, the average rates for the repair (μ) of the equipment contained in the referred station are defined, i.e., the Mean Time to Repair (MTTR) for each station, Equation (31).
3. The Mean Time to Failure (λ) (MTTF) of each equipment is defined, as in Equation (30).
4. The productivity ($\mathcal{Q}$) of each machine assigned to the corresponding workstation. It is defined as the relationship between the volume of output and the volume of inputs. This information is provided by the manufacturing engineering area and is obtained from historical data.
5. The efficiency (η) of the machine installed on the corresponding workstation is defined. Mechanical efficiency is a measure of how well the machine converts the input work or energy into some useful output. In our case, we use Equation (34) to calculate this parameter.

6. The analysis provides the expected productivity of each machine ($\overline{\mathcal{Q}}$) associated with a workstation, Equation (35).
7. Lines 6 through 12 define the expected productivity per station $(\overline{\mathcal{Q}}_{kj}) = f_{kj}\,\overline{\mathcal{Q}}_j$.
8. Row 18 shows the accumulated values of the expected production at station j. That is, the total average parts manufactured at station j, i.e.,

$$\text{Average productivity} = \sum\nolimits_{j=6}^{17} \overline{\mathcal{Q}}_{kj},\ \ i = 1, \ldots, 12 \tag{38}$$

9. Finally, the leisure index ($\mathcal{L}$) generated in each station, due to its unavailability, is also presented as in Equation (37).

Some Additional Comments on the Results

Table 7 fully summarizes the results in this research. Some important aspects to highlight are the following. This methodology was applied to a real metalworking manufacturing model in a refrigerator manufacturing company in Mexico. The actual results were flattering as they are fairly close to the values obtained onsite. It is important to note that the entire project is based on the knowledge of the matrix of transition probabilities. Likewise, the knowledge of the availability of the equipment is essential for the calculation of various parameters.

It is also important to note that at the start of system operations, it is verified that

$$\boldsymbol{P^t}\,\boldsymbol{\pi^0} \to \boldsymbol{\pi}\ \in \mathbb{R}^{12}$$

as $t \to \infty$.

Under stable state conditions, it can be assumed that each component of π is equiprobable, then it is verified that $\lim_{t\to\infty} \boldsymbol{\pi}\,\boldsymbol{P^t} \overset{\mathfrak{D}}{\overbrace{\to}} \boldsymbol{\pi^t}$, where the symbol $\overset{\mathfrak{D}}{\overbrace{\to}}$ means convergence in distribution.

Finally, the approximation of the duration of visits to each station can be approximated using Equation (15) as long as the system is in steady state conditions.

$$\mathbb{P}\left(L_i^t = k \,\middle|\, \tau_i^{n-1} < \infty\right) = P\,[\,\mathrm{X}(t)] = k. \tag{39}$$

It is interesting to highlight some efforts made to model mass manufacturing systems with the approach proposed here, see, for example, [26–30]. This proposal is a first attempt to develop a model in process manufacturing systems, and more research is needed in the future to improve the modeling.

6. Conclusions

In this proposal Markov chain theory was used to model the activity of a manufacturing line by process. Due to the frequency with which these types of layouts are found in manufacturing engineering, the problem addressed here acquires special importance in order to determine the average production that can be aspired to when the failure and maintenance rates of the system are known.

Other interesting lines of research to explore, in the future, are the association of costs to the process, the time spent at each station (Equation (14) provides an estimator of this), the rejected materials and the presence of preceding buffers. In practice, manufacturing engineers face these challenges using a Monte Carlo simulation in order to obtain the mentioned estimators. However, the credibility of such models depends largely on the quality of the proposed model and the ability to program it. Therefore, an approximate solution is always preferable, despite its uncertainty.

The real results obtained in this project showed its feasibility and above all, its usefulness in terms of obtaining quantitative indicators of the expected production in a manufacturing line with these characteristics. This offers advantages over simulation since the estimators depend only on the matrix of transition probabilities.

Finally, it is important to point out that the reliability of the chain run can be easily obtained from the product of the reliability of each station.

Author Contributions: G.P.-L. developed the mathematical model of the proposal and obtained the information in the company where it was applied. He also worked with the data and programmed the code. J.F.M.-S. wrote the document and gave statistical treatment to the sample information obtained in situ for the preparation of the document. F.V.-M. adapted the mathematical model to the information obtained from the company. He also adapted some concepts of the mathematical model. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data are available upon request.

Acknowledgments: The authors thank the reviewers.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

FMS	Flexible Manufacturing System
MPS	Master Production Schedule
(s, S)	A minimum/maximum inventory policy
CONWIP	Constant Work in Process
KANBAN	Japanese word meaning sign
MRP	Material Requirements Planning
MRP II	Manufacturing Resource Planning
MPS	Master Production Schedule

References

1. Lechuga, G.P.; Sánchez., F.M. Modeling and Optimization of Flexible Manufacturing Systems: A Stochastic Approach. In *Intelligent Computing & Optimization, Proceedings of the ICO 2018 International Conference on Intelligent Computing & Optimization, Pattaya, Thailand, 4–5 October 2018*; Advances in Intelligent Systems and Computing; Springer: New York, NY, USA, 2018; pp. 539–546. Available online: https://link.springer.com/chapter/10.1007/978-3-030-00979-3_57 (accessed on 24 October 2021).
2. Althobaiti, O.S. IoT Limitations and Concerns Relative to 5G Architecture. In *Handbook of Research on Smart Technology Models for Business and Industry*; IGI-Global: Hershey, PA, USA, 2020. Available online: https://www.igi-global.com/chapter/iot-limitations-and-concerns-relative-to-5g-architecture/259125 (accessed on 14 September 2021). [CrossRef]
3. Lechuga, G.P.; Gress, E.S.H.; Karelyn, A.A.; Orán, M.G.M. Some Efficiency Measures in the Operation of Flexible Manufacturing Systems: A Stochastic Approach, Society for Industrial and Applied Mathematics. In Proceedings of the 2009 SIAM Conference on Mathematics for Industry, San Francisco, CA, USA, 9–10 October 2009; Society for the Industrial and Applied Mathematics (SIAM): San Francisco, CA, USA, 2009; pp. 31–36. [CrossRef]
4. Lechuga, G.P.; Karelyn, A.; Oran, M.M.; Aragones, L.P. Stochastic analysis of operations decoupling in flexible manufacturing systems. In Proceedings of the 5th WSEAS/IASME International Conference on Systems Theory and Scientific Computation, Malta, 15–17 September 2005; pp. 169–174. Available online: https://dl.acm.org/doi/10.5555/1373616.1373645 (accessed on 20 October 2021).
5. Nor, A.R.; Sharif, S.M.; Esa, M.M. Lean Manufacturing Case Study with Kanban System Implementation, International Conference on Economics and Business Research 2013 (ICEBR 2013). *Procedia Econ. Financ.* **2013**, *8*, 174–180. [CrossRef]
6. Lechuga, G.P.; Venegas-Martínez, F.; Ramírez, P. Stochastic optimization of manufacture systems by using Markov decision processes. In *Handbook of Research on Modern Optimization Algorithms and Applications in Engineering and Economics*; Vasant, P., Weeber, G.-W., Dieu, V.N., Eds.; IGI-Global: Hershey, PA, USA, 2016; pp. 185–202. Available online: https://books.google.com.mx/books?hl=es&lr=&id=HvS1CwAAQBAJ&oi=fnd&pg=PP1&dq=Handbook+of+Research+on+Modern+Optimization+Algorithms+and+Applications+in+Engineering+and+Economics&ots=A8JegKgE4q&sig=VCwtCOZBhSGnW5QmQCxB9Ys44M8#v=onepage&q=Handbook%20of%20Research%20on%20Modern%20Optimization%20Algorithms%20and%20Applications%20in%20Engineering%20and%20Economics&f=false (accessed on 27 September 2021).
7. Buzacott, J.; Shanthikumar, G. *Stochastic Models of Manufacturing Systems*; Prentice Hall, Inc.: Hoboken, NJ, USA, 1993.
8. Gershwin, S.B. *Manufacturing Systems Engineering*; Prentice Hall: Hoboken, NJ, USA, 1994.

9. Shanthikumar, J.G.; Yao, D.D.; Zijm, W.H.M. *Stochastic Modeling and Optimization of Manufacturing Systems and Supply Chains*; International Series in Operations Research & Management Science; Springer: Berlin/Heidelberg, Germany, 2003; Volume 63. Available online: https://www.amazon.com/Stochastic-Optimization-Manufacturing-International-Operations/dp/1402075081 (accessed on 2 September 2021).
10. Suresh, P.S.; Zhang, H.Q.; Zhang, Q. *Average-Cost Control of Stochastic Manufacturing Systems*; Stochastic Modelling and Applied Probability; Springer: Berlin/Heidelberg, Germany, 2005; Volume 54. Available online: https://www.amazon.com/-/es/Suresh-P-Sethi/dp/0387219471 (accessed on 12 September 2021).
11. Liberopoulos, G.; Papadopoulos, C.T.; Tan, B.; Smith, J.M.; Gershwin, S.B. (Eds.) *Stochastic Modeling of Manufacturing Systems, Advances in Design, Performance Evaluation, and Control Issues*; Springer: Berlin/Heidelberg, Germany, 2016.
12. Van Vuuren, M.; Adan, I.J.; Resing-Sassen, A. Performance Analysis of Multi-Server Tandem Queues with Finite Buffers and Blocking. *OR Spectr.* **2005**, *27*, 315–338. Available online: https://www.yumpu.com/en/document/view/36573212/performance-analysis-of-multi-server-tandem-queues-with-cqm (accessed on 30 September 2021). [CrossRef]
13. Stefansson, H.; Sigmarsdottior, S.; Jensson, P.; Shah, N. Discrete and Continuous Time Representations and Mathematical Models for Large Production Scheduling Problems: A Case Study from the Pharmaceutical Industry. *Eur. J. Oper. Res.* **2011**, *215*, 383–392. Available online: https://www.infona.pl/resource/bwmeta1.element.elsevier-3028e374-5eec-39c8-a4a2-00ccc864ab2b (accessed on 30 September 2021). [CrossRef]
14. Alfieri, A.; Matta, A. Mathematical programming formulations for approximate simulations of multistage production systems. *Eur. J. Oper. Res.* **2012**, *219*, 773–783. [CrossRef]
15. Stroiteleva, T.; Vukovich, G. Mathematical Modeling of Workflows in Production Systems. *Mod. Appl. Sci.* **2015**, *9*, 173–183. Available online: https://www.researchgate.net/publication/276576036_Mathematical_Modeling_of_Workflows_in_Production_Systems (accessed on 2 September 2021).
16. Tan, B. Mathematical programming representations of the dynamics of continuous-flow production systems. *IIE Trans.* **2015**, *47*, 173–189. [CrossRef]
17. Buyukdagli, O.; Fadıloğlu, M.M. A Mathematical Programming Framework for Control of Tandem Production Lines. In Proceedings of the SMMSO 11th Conference on Stochastic Models of Manufacturing and Services Operations, Lecce, Italy, 4–9 June 2017. [CrossRef]
18. Zhang, B.; Yang, Y. Capacity assignment of the production lines with consideration of equilibrium requirement in cold rolling area. In Proceedings of the 27th Chinese Control and Decision Conference (2015 CCDC), Qingdao, China, 23–25 May 2015; pp. 1662–1665. [CrossRef]
19. Roshani, A.; Giglio, D. A Mathematical Programming Formulation for Cost-Oriented Multi-Manned Assembly Line Balancing Problem. *IFAC-PapersOnLine* **2015**, *48*, 2293–2298. [CrossRef]
20. Black, A.J.; McKane, A.J. Stochastic formulation of ecological models and their applications. *Trends Ecol. Evol.* **2012**, *27*, 337–345. [CrossRef] [PubMed]
21. Qi, W.; Zong, G.; Karimi, H.R. L_∞ Control for Positive Delay Systems with Semi-Markov Process and Application to a Communication Network Model. *IEEE Trans. Ind. Electron.* **2019**, *66*, 2081–2091. [CrossRef]
22. Zhang, L.; Prieur, C. Stochastic stability of Markov jump hyperbolic systems, with application to traffic flow control. *Automatica* **2017**, *86*, 29–37. [CrossRef]
23. Smith, J.; MacGregor, B.; Tan, B. (Eds.) *Handbook of Stochastic Models and Analysis of Manufacturing System Operations*; International Series in Operations Research & Management Science; Springer: Berlin/Heidelberg, Germany, 2013; Available online: https://www.springer.com/gp/book/9781461467762 (accessed on 12 September 2021).
24. Parzen, E. Stochastic Processes, Classics in Applied Mathematics, Society for Industrial and Applied Mathematics (SIAM), cl24. 1999. Available online: https://epubs.siam.org/doi/book/10.1137/1.9781611971125?mobileUi=0 (accessed on 17 September 2021).
25. Latouche, G.; Ramaswami, V. *Introduction to Matrix Analytic Methods in Stochastic Modeling*; Society for Industrial and Applied Mathematics (SIAM): Philadelphia, PA, USA, 1999.
26. Leigh, J.M.; Jackson, L.; Dunnett, S.; Lugo, H.; Sharpe, R.; Neal, A.; West, A. *Modelling Manufacturing Processes Using Markov Chains*; Conference Contribution; Loughborough University: Loughborough, UK, 2017; Available online: https://hdl.handle.net/2134/26018 (accessed on 20 September 2021).
27. Scrivano, S.; Tolio, T. A Markov Chain Model for the Performance Evaluation of Manufacturing Lines with General Processing Times. *Procedia CIRP* **2021**, *103*, 20–25. [CrossRef]
28. Hadjinicola, G. A Markov Chain Model of Serial Production Systems with Rework, Hadjinicola 2002. 2002. Available online: https://www.semanticscholar.org/paper/A-MARKOV-CHAIN-MODEL-OF-SERIAL-PRODUCTION-SYSTEMS-Hadjinicola/afcdcc1d564e6a6c82d9a4b606c9ad481902e787 (accessed on 14 October 2021).
29. Ramaraj, G. Production Planning in Different Stages of a Manufacturing Supply Chain under Multiple Uncertainties. Master's Thesis, Iowa State University, Ames, IA, USA, 2017. Available online: https://lib.dr.iastate.edu/etd/16108 (accessed on 27 October 2021).
30. Madhusudanan, P.V.; Chandrasekharan, M.P. An absorbing Markov chain model for production systems with rework and scrapping. *Comput. Ind. Eng.* **2008**, *55*, 695–706. [CrossRef]

mathematics

MDPI

Article

The Interplay between Digital Entrepreneurship and Sustainable Development in the Context of the EU Digital Economy: A Multivariate Analysis

Emilia Herman

Faculty of Economics and Law, George Emil Palade University of Medicine, Pharmacy, Science and Technology of Targu Mures, 540139 Targu Mures, Romania; emilia.herman@umfst.ro

Abstract: A real challenge for the EU economy and society is to achieve both green and digital transitions in order to tackle the major economic, social and environmental issues faced by EU member states. In this context, digital entrepreneurship, which lies at the intersection of digital technologies and entrepreneurship, has recently benefited from increasing attention both in theoretical and empirical research and in strategic policies. Given these aspects, the aim of this article was to investigate the interrelationship between digital entrepreneurship and productive and innovative entrepreneurship and its impact on the achievement of the Sustainable Development Goals (SDGs) in EU countries. The results of correlation and regression analysis revealed that digital entrepreneurship, which implies productive and innovative entrepreneurial activities, is positively influenced by the degree of a country's digitalization and, in turn, has a positive impact on the achievement of the SDGs (Total SDGs, SDG 8 and SDG 9). Furthermore, the findings of the principal component analysis and cluster analysis emphasize that there are differences and common features between EU countries in terms of the interrelationship between digital entrepreneurship, digitalization, economic development, national competitiveness and achievement of the SDGs. Therefore, specific measures should be implemented to boost digital entrepreneurship (especially in some central and eastern EU countries) so that this will be the key driver for sustainable development.

Keywords: digital entrepreneurship; productive entrepreneurship; digitalization; sustainable development; SDGs; EU countries; regression analysis; principal component analysis; cluster analysis

MSC: 62H25; 62J05

1. Introduction

The "Transforming our World: the 2030 Agenda for Sustainable Development" resolution adopted by the United Nations (UN) General Assembly [1] in 2015 as a comprehensive policy blueprint sets 17 Sustainable Development Goals (SDGs), which represent the global priorities for development by 2030 and address the major economic, social and environmental challenges faced by global and national communities. The UN 2030 Agenda attempts to restore harmony between progress and sustainability by creating a sustainable world which includes all countries [2] and achieve the kind of "development that meets the needs of the present without compromising the ability of future generations to meet their own needs" [3] (p. 43). At the European Union (EU) level, all SDGs are an intrinsic part of EU strategies and are regarded as essential for the formulation of policies either for internal and external action in all sectors [4]. Moreover, the European Commission [4] recognizes that the full accomplishment of the UN 2030 Agenda plays a pivotal role in boosting resilience and making sure that economies can recover from future shocks as the world faces the twin green and digital transitions.

Entrepreneurship and innovation have been recognized by the UN [5] as key drivers for sustainable development, addressing all three challenges of sustainability (economic, social and environmental) in the context of the 2030 Agenda. Researchers around the world have highlighted the role of entrepreneurship in sustainable development by focusing more on the economic and social dimensions of sustainability. Thus, entrepreneurship is seen as a core factor in generating wealth [6], driving economic growth and development, creating jobs and reducing unemployment and poverty, improving well-being and people's standard of living [7–11]. Furthermore, empirical research [12–15] proves that a high level of quality of entrepreneurship (productive, innovative and opportunity-driven entrepreneurship) is more important than a high quantity of entrepreneurship for producing a positive effect on sustainable development. Even though the entrepreneurship literature has increasingly recognized entrepreneurship as an effective solution to various social, economic and environmental challenges of sustainable development [10,16,17], a research gap has been identified in terms of the holistic approach to the simultaneous role of the entrepreneurial activity that could address all three challenges [18,19].

Over the last decade, on the Fourth Industrial Revolution's background, entrepreneurship has been significantly transformed due to the influence of digitalization on the economy and society [20–23]. Therefore, in recent years, digital entrepreneurship, which according to the EC has made use of novel digital technologies to shape existing businesses and to impact new ones [24], has received increasing attention both in theoretical and empirical research [7,25,26].

The interplay between digitalization and sustainability is highly emphasized by several research studies [27–29], which recognize that digitalization can be the right solution for research gaps and societal problems, while sustainability is a prerequisite for a responsible digital transformation [2]. Digitalization may be defined as the "sociotechnical process of applying digitizing techniques to broader social and institutional contexts that render digital technologies infrastructural" [30] (p. 749), as well as "the process by which economy and society are evolved around digital technologies providing new added-value opportunities" [2] (p. 15).

Deep integration of digital technologies in the economy has a high potential to contribute to sustainable development [31]. Empirical studies [32,33] highlight that the success of digital transformation of businesses in any sector of the economy (from the manufacturing sector to the education sector) depends on multiple factors, such as external support for digitalization, digitalization readiness preassessment and information and digital technology readiness [32]. In the current environment, one of the biggest challenges that any business or society faces, identified by Jafari-Sadeghi et al. [20], is the way in which new digital technologies are adopted, integrated and exploited.

As one of the connecting mechanisms between the multiple dimensions (technological, economic, social and ecological) of the socioeconomic system, digitalization can be a serious source of challenges to the resilience of this system [34,35], providing both new opportunities and new risks with unpredictable consequences. Such challenges need to be managed in a sustainable way, relying on the principles of sustainable development [35]. Moreover, at the EU level, the new growth strategies (e.g., The European Green Deal (2020) and Shaping Europe's Digital Future (2020)) acknowledge the need for a holistic approach to the nexus between digitalization and sustainability to achieve the twin EU transitions. Therefore, emphasis is placed on that approach to investigate the multiple roles of digital entrepreneurship in the economic, social and economic dimensions of sustainable development.

Bearing this context in mind, *the aim of this article* is to highlight the interrelationship between digital entrepreneurship and productive and innovative entrepreneurship and its impact on the achievement of the SDGs (Total SDGs, SDG 8 and SDG 9) in the EU member states for the 2018–2019 period. Additionally, this research highlights the specific measures that should be taken to improve digital entrepreneurship in order to increase its effect on the SDGs. The *research objectives* focused on (1) analyzing the link between productive and innovative entrepreneurship on the one hand and economic development, national competitiveness and the degree of digitalization of an economy and society on the other,

(2) exploring the interlink between digital entrepreneurship, productive entrepreneurship and innovative entrepreneurship, (3) investigating the nexus between digital entrepreneurships and the degree of digitalization of an economy and society, (4) analyzing the impact of digital entrepreneurship on the achievement of the SDGs, and (5) identifying the differences and common features between EU countries based on their interrelationship between digital, innovative and productive entrepreneurship, digitalization, economic development, national competitiveness and achievement of the SDGs.

Although several studies address the influence of digital technologies and/or entrepreneurship on achieving sustainable development, there is a lack of empirical research on the effects of digital entrepreneurship on sustainable development. In this research paper, as opposed to the existing empirical studies, we investigate the impact of digital entrepreneurship as the combination of digital technologies and entrepreneurship on the simultaneous attainment of the economic, social and environmental goals of sustainable development. Thus, this research contributes to filling the gap of empirical studies that highlighted the impact of digital entrepreneurship on sustainable development through a holistic and integrative approach.

The next section discusses the theoretical framework and the related literature, as well as the research hypotheses. Section three outlines the empirical data and statistical methods used. The fourth section provides the results of the research, the discussion and their main implications. The last section presents the article's conclusions as well as its limitations and the further possible research.

2. Theoretical Background and Research Hypotheses

2.1. The Impact of Productive and Innovative Entrepreneurship on the Economic Dimension of Sustainable Development

Scholars have acknowledged a major impact of entrepreneurship on sustainable development, but they tend to focus more on the economic dimension of sustainability. Thus, entrepreneurship and its potential impact on macroeconomic performance have been largely studied over the last two decades [10,36,37], and both topics still continue to be of interest. Most of the empirical studies highlighted that entrepreneurship positively impacts the national economy in terms of economic growth, economic development and national competitiveness [7,11,15,38,39], and they also underlined that the degree of this impact in turn depends on a variety of influencing factors. At the same time, there are other empirical studies [10,17,40] which concluded that this macroeconomic impact of entrepreneurship is negative. These contradictory results can have multiple and interrelated explanations, such as the complexity of the link between entrepreneurship and economic growth [9], the level of economic development, the country's development stage [12], the country's innovativeness [38,41], the institutional and cultural settings [38], the motivations of entrepreneurs [8,42], the definition and measurement of entrepreneurship [15,36,40] and different forms of entrepreneurship [43,44].

As empirical research [14,36,40] has shown, the influence of entrepreneurship on economic welfare depends on how it is defined and measured. Thus, entrepreneurship measured by indicators which reflect a high level of innovativeness and quality (e.g., the Global Entrepreneurship Index, innovative SMEs, innovative and high-growth firms and opportunity-driven entrepreneurship) has a positive macroeconomic impact [8,13,15], and this underlines that the quality of entrepreneurship matters more than the quantity. In the same vein, Szerb et al. [15], analysing the relationship between entrepreneurship and regional performance at the level of the 121 EU regions in the 2012–2014 period, pointed out that the regional performance was negatively impacted by quantity entrepreneurship, while in the case of quality (Schumpeterian) entrepreneurship, this impact was positive. Opportunity-driven entrepreneurship is positively related to job growth and economic growth and development [8,12,38], as well as to technological and innovation progress [45]. On the contrary, it was found that necessity entrepreneurship is inversely linked to country-level innovation [45] and economic development [38]. Moreover, Dhahri et al. [46], examining the impact of opportunity and necessity-driven entrepreneurship on all three dimensions of sustainable development based on data for 20 developing countries, highlighted the positive impact of opportunity-driven entrepreneurship on all three dimensions of sustainable development (SD). Thus, it is claimed that opportunity entrepreneurs have a key role in achieving the SDGs. It is also shown that necessity-driven entrepreneurship only has a negative effect on the environmental sustainability dimension.

According to Baumol [47], entrepreneurship can be "productive" (contributing to an economy's productivity growth), "unproductive" and even "destructive" due to the division of entrepreneurial activities into productive (based on innovation) and unproductive activities. This division is mainly determined by national institutional settings (rule of game) which can modify the incentive structures that prompt people to choose productive entrepreneurship, which fosters growth [48]. The systematic literature review conducted by Urbano et al. [37] claims that formal as well as informal institutions influence the nexus between entrepreneurship and economic growth. Moreover, from a holistic perspective, entrepreneurial activity emerges from the interaction between institutions, stakeholders and entrepreneurs themselves [25,49,50]. In the same light, Cao and Shi [51] found that structural gaps, resource scarcities and institutional voids are some of the main challenges countries face in adoption of the entrepreneurial ecosystems model from advanced economies to emerging economies. In the systemic and integrative approach, the *entrepreneurship ecosystem* is defined by the researchers of the Global Entrepreneurship and Development Institute as "the dynamic, institutionally embedded interaction between entrepreneurial attitudes, entrepreneurial abilities, and entrepreneurial aspirations by individuals, which drives the allocation of resources through the creation and operation of new ventures" [52] (p. 15). Based on this definition, the same researchers measured the entrepreneurship ecosystem through the Global Entrepreneurship Index (GEI), which reflects both opportunity-driven entrepreneurship and productive entrepreneurship that generates wealth [52].

Taking into consideration the above-mentioned perspectives, in the current paper, the *Global Entrepreneurship Index* was used to analyze the interlink between productive entrepreneurship and sustainable development.

Innovative entrepreneurship is considered one of the most "productive" forms of entrepreneurship [43,47]. Significant economic growth and a high level of economic development and national competitiveness are mostly generated by a small number of new firms or ventures, especially *innovative new firms* (innovative SMEs) and enterprises with high growth expectations [10,15,36,38]. Entrepreneurship, defined as a process of opportunity recognition and exploitation as well as a process of creation of new goods and services, is inseparable from innovation [16]. Thus, both innovation and entrepreneurship are two complementary concepts which are positively linked to each other [36,53,54]. Moreover, both concepts have been intertwined with economic development, as stated by Schumpeter [55]. Innovative entrepreneurs, according to Schumpeter [55], recognize and exploit opportunities, providing new combinations (innovations) as new methods of production and organization, new products, new markets, etc. Therefore, through innovation, the old structure is incessantly destroyed, and a new one is created ("creative destruction"), which means the creation of new ways to meet demands.

Entrepreneurs make a substantial contribution to growth, productivity and social welfare by using their knowledge and putting forth revolutionary innovations that serve our society [48]. Innovation and entrepreneurship are key factors of national competitiveness [56–58], alongside other influencing factors such as digitalization, institutions, infrastructure, information and communication technology (ICT) adoption, health and education [57,59]. Pradhan et al. [9] investigated the short- and long-term impacts of innovation and entrepreneurship on economic growth in the case of Eurozone countries and found that these three variables are intricately intertwined. In a sample of 64 countries worldwide (factor-driven, efficiency-driven and innovation-driven economies), Du and O'Connor [14] showed that improvement-driven opportunity entrepreneurship significantly contributes to increasing the national level of efficiency (expressed by GDP/labor). Ivanović-Đukić et al. [43], investigating the effect of various forms of innovative entrepreneurship (new technology development entrepreneurship, new product entrepreneurship and high growth expectation entrepreneurship) in 21 EU countries, found that all types of innovative entrepreneurship have positively influenced economic growth, but the effect is different depending on the *degree of digitalization of a country*. Digitalization is seen as a source of innovation on the one hand and, on the other, an outcome by restructuring business patterns in all sectors of economy, while the entrepreneurs are both the drivers and the affected agents of digital transformations [35].

Based on these premises, we developed the following hypotheses:

Hypothesis 1. *In the EU countries with higher economic development and national competitiveness, productive and innovative entrepreneurship is higher.*

Hypothesis 2. *Productive and innovative entrepreneurship is higher in the EU countries with a high degree of digitalization of the economy and society.*

2.2. From Productive and Innovative Entrepreneurship to Digital Entrepreneurship

Although digitalization affects more and more aspects of our lives, it mainly determines the transformation of the entrepreneurial process and business models in various sectors as a response to the change in a society's needs [35]. For businesses, the digital transformation, induced by increased use of digital technologies, contributes to changes in products or services, business models and the way in which products or services are manufactured and delivered [20,60,61]. According to the European Commission [24], the distinguishing feature of *digital enterprises* is their high digital intensity (using new digital technologies such as mobile and cloud solutions, big data analytics and social media), which improves business operations, generates new business models and creates growth and jobs. Enterprises can benefit from the adoption of digital technologies in multiple ways, such as lower operational costs, higher annual turnover and productivity, more competitive advantages and new business model opportunities [7,60–64]. Martin-Rojas et al. [65] pointed out that corporate entrepreneurship is positively influenced by the acquisition and integration of technology and infrastructure, and in turn, it positively influences organizational performance in terms of profitability and growth. Moreover, researchers highlight that digital technologies generate a myriad of sources of opportunities for entrepreneurs to create a new generation of start-ups or new ventures [21,62,63,66], as well as a range of challenges to business owners and their firms [7]. These new business ventures and digital start-ups make novel technologies an integral part of their business models and operations [67]. Nambisan et al. [68] emphasized that both open innovation, which implies more open and distributed innovation models, and increasing digital platforms as a path to value generation have changed the nature of entrepreneurship throughout industries and have generated new opportunities for entrepreneurs and their firms.

As Jafari-Sadeghi et al. [20] underlined, over the last decade, entrepreneurship has been significantly transformed due to the interaction between digital technologies, platforms and infrastructures and its influence on value creation. The same empirical study proved that the technology readiness factors (ICT investment and ICT usage by businesses) had a positive impact on technology-driven entrepreneurship in European countries in the 2009–2015 period.

Taking into account that new digital technologies have transformed the entrepreneurial processes and outcomes, *digital entrepreneurship is the result of the intersection of digital technologies and entrepreneurship* [23,66]. Studies [66,67] highlighted that the core of digital entrepreneurship is the identification and pursuit of entrepreneurial opportunities presented by digital technologies. Digital entrepreneurship has not been clearly defined yet, and according to Kraus et al. [25], the body of research on digital entrepreneurship is still in a developmental stage.

In the current study, according to the European Commission's definition [24] of digital entrepreneurship, and in line with other authors [25,69–71], digital entrepreneurship is seen as a subcategory of entrepreneurship which entails the digitalization of some or all physical aspects of a traditional business. It can also be seen "as the reconciliation of traditional entrepreneurship with the new way of creating and doing business in the digital era" [72] (p. 1). Taking into consideration that digital entrepreneurship is a multilevel phenomenon [7,63,66], theoretical and empirical approaches to digital entrepreneurship have spread across multiple levels, from characteristics of digital entrepreneurs and digital firms to communities as a whole ecosystem [21,23,66,73]. Researchers have shown that the success of a digital start-up depends on the characteristics of the entrepreneurs and internal operations of an enterprise, as well as the characteristics of the community to which it belongs [21,63]. Moreover, the integration of digital technologies in the business process does not rely solely on the internal changes of entrepreneurial processes and organizational management but also on external system conditions (institutional role, digital infrastructure and digital marketplace tendencies) and social attitudes (digital skills, digital trust, technology adoption, etc.) [35]. Other studies [23,62] revealed that digital entrepreneurship distinguishes itself by less bounded and more networked processes and outcomes which vary according to space and time [7]. Therefore, recent studies [66,67,74] focused on the digital entrepreneurship ecosystem which takes into account a multilevel perspective of digital entrepreneurship and captures the interaction between a larger number of actors (entrepreneurs, stakeholders, institutions, etc.). In the context of the current digital era, Autio et al. [21] suggest that an entrepreneurial ecosystem is a digital economy that utilizes the digital technologies and infrastructures available to make the pursuit of entrepreneurial opportunities easier with the help of new ventures by rethinking the business model. As Susan and Acs [44] showed, the digital entrepreneurial ecosystem lies at the intersection of the entrepreneurial ecosystem (agents (entrepreneurs) and institutions) and digital ecosystem (users and digital infrastructure). Autio et al. [26], in order to elaborate upon the *European Index of Digital Entrepreneurship Systems (EIDES)*, took into account that an entrepreneurial ecosystem is a community of different stakeholders such as entrepreneurs, accelerators and advisors that use specialized resources to support entrepreneurial stand-up, start-up and scale-up through digitally upgraded business models. Empirical evidence at the EU level (27 EU countries and the United Kingdom) shows that the EIDES is positively correlated with economic development (GDP/capita) and national competitiveness (Global Competitiveness Index) [26]. In the current study, we take into consideration this perspective of the digital entrepreneurial ecosystem, and we use the EIDES to analyze digital entrepreneurship and its relationship with sustainable development.

This study explores how digital entrepreneurship is interlinked with productive and innovative entrepreneurship and investigates to what extent digital entrepreneurship can be influenced by the degree of digitalization of an economy and society in the case of EU countries. Thus, we developed the following hypotheses:

Hypothesis 3. *Digital entrepreneurship interrelates with productive and innovative entrepreneurship in EU countries.*

Hypothesis 4. *The degree of digitalization of an economy and society in EU countries positively influences digital entrepreneurship.*

2.3. Digital, Productive and Innovative Entrepreneurship for the SDGs' Achievement

Dhahri and Omri [18] emphasized that there is a research gap in the holistic approach which assesses the impact of the entrepreneurial activity on the simultaneous reaching of the economic, social and environmental goals of sustainable development. Thus, the impact of entrepreneurship on the social and environmental dimensions of sustainable development is less known and researched than its economic impact [12]. Several researchers [6,19,75,76] recognized entrepreneurship as an efficient solution to environmental degradation and social inequality more than a potential cause of them. Moreover, empirical studies [77,78] have analyzed economic growth and its impact on poverty reduction through entrepreneurship, innovation and the development of new technology. In the same vein, Si et al. [77] stated that in recent years, inclusive entrepreneurship which incorporates more and more digital technology is seen as an effective solution to the alleviation of poverty and thus the reduction of social inequality.

Another empirical study [18], based on a sample of 20 developing countries in the 2001–2012 period, claims a positive contribution to the economic and social dimensions of sustainable development but a negative effect of entrepreneurship on the environmental dimension. Youssef et al. [79] found that entrepreneurship can contribute to all dimensions of sustainable development but only under specific circumstances related to innovation and institutional quality. Thus, despite the positive effect on economic growth, the authors proved, in the case of 17 African countries, that formal and informal entrepreneurship contribute to environmental degradation (CO_2 emissions), which is much higher in the case of informal entrepreneurship. The same authors [79] also proved that high levels of innovation and institutional quality enhance the strongly positive impacts of (formal and informal) entrepreneurship on all dimensions of sustainability.

In line with this, according to the Global Entrepreneurship Monitor Report [12], entrepreneurial activity is a key driver of the achievement of multiple SDGs, such as SDG 1 ("End poverty in all its forms everywhere"), SDG 8 ("Promote inclusive and sustainable economic growth, employment and decent work for all") and SDG 10 ("Reduce inequality within and among countries"). Moreover, the UN [5] has recognized the key role of entrepreneurship in achieving all three dimensions of SD. Thus, the UN [5] has acknowledged entrepreneurship and innovation as key drivers to harness the national economic potential for achieving the SDGs. Regarding the contribution to the *economic dimension* of sustainability, the UN recognizes that entrepreneurship generates economic growth and creates jobs, fostering decent work and innovation. As for the *social dimension* of sustainability, it has been recognized that entrepreneurship can positively contribute to reducing inequalities, stimulate social cohesion and enlarge opportunities for all [5]. The contribution of entrepreneurship to the *environmental dimension* is also recognized by the UN [5], which highlights that entrepreneurship can respond to environmental challenges through promoting eco-friendly consumption patterns and sustainable practices, as well as through advocating for the implementation of novel digital technologies and resilience policies.

Although entrepreneurship is directly or indirectly related to all 17 SDGs, we focused in this paper on SDG 8 and SDG 9, the goals which highlight the interlink between economic growth, entrepreneurship, enterprises and innovation technology and its impact on sustainable development. Thus, *SDG 8* ("Promote sustained, inclusive and sustainable economic growth, full and productive employment and decent work for all") emphasizes the importance of sustained economic growth in order to generate well-paid jobs and decent work, which improve the quality of life as well as eco-efficiency through resource efficiency in consumption and production [1,6]. The EU Report [80] underlines that the prosperity of European countries as well as the well-being of individuals are ensured by inclusive green economic growth and decent and productive employment. Moreover, *Target 8.3 of SDG 8* reflects the importance of the closed interrelation between entrepreneurship and innovation in order to ensure both green and sustainable economic growth for all [1]. *SDG 9* ("Build resilient infrastructure, promote inclusive and sustainable industrialization and foster innovation") posits resilient and sustainable infrastructure, inclusive and sustainable industrialization and research and innovation as solutions to economic, social and environmental challenges [1]. In addition, this goal emphasizes the major role that innovation and technology play in reaching the SDGs. In this context, entrepreneurship, as both the engine of technological innovation and the answer to it, can generate economic growth, labor productivity and income, competitiveness, resource efficiency and job creation, as well as improve health, education and well-being for all [4,6,80].

Moreover, recent studies and national and international reports [2,27,28,81,82] acknowledged the essential role of technological innovation and digital technology in achieving every one of the 17 SDGs. For example, based on a sample of 75 low-, middle- and high-income countries, Omri [81] found that the effect of technological innovation on the Sustainable Development Goals depends on the stages of economic development. Thus, technological innovation simultaneously contributes to the economic, environmental and social dimensions of sustainable development only in high-income countries. In the case of middle-income countries, technological innovation only influences the economic and environmental dimensions, and in the case of low-income countries, no effects on these dimensions were identified. Several studies [27,70,83] have highlighted that the optimization of resources through the adequate adoption and use of digital technology can be enhanced, and businesses and entrepreneurship can become more sustainable. Furthermore, they pointed out that the best solution to achieve the SDGs is (digital) entrepreneurship, based on the principles of cutting-edge technologies, open innovation and social business.

Thus, although several empirical studies explore the role of digital technologies and entrepreneurship in achieving sustainable development, there is a lack of empirical research on the effects of digital entrepreneurship on the SDGs. Therefore, this study attempts to fill this research gap by exploring to what extent digital entrepreneurship can be a key factor for the achievement of the SDGs in the EU member states. Although this paper focuses more on the impact of digital entrepreneurship on SDG 8 and SDG 9, it offers an overall picture of the extent to which each country fulfils all the SDGs. As well as assessing the impact of digital entrepreneurship on the simultaneous attainment of the economic, social and environmental goals of sustainable development, we also take into consideration the total SDGs as the result of the average of all 17 SDGs. Given these aspects, we developed the following hypotheses:

Hypothesis 5. *Digital entrepreneurship has a positive impact on the achievement of the SDGs (Total SDGs, SDG 8 and SDG 9).*

Hypothesis 6. *There are differences and common features between the EU countries based on their interrelations between (digital, innovative and productive) entrepreneurship, digitalization, economic development, national competitiveness and achievement of the SDGs (Total SDGs, SDG 8 and SDG 9).*

3. Research Data and Methods

3.1. Data and Sample

In order to test the research hypotheses and thus achieve the aim of this study, we analyzed the variables described in Table 1. The data analysis was performed for the 2018–2019 period. The limitation of the analysis to this period was determined by the fact that some composite indexes (EIDES, SDGs Index, GCI 4.0. and NRI) were calculated only for the last 2 or 3 years. Because of time lags in data creation and release, the impact of the COVID-19 pandemic was not identified.

Our sample consisted of 25 countries from the EU without Luxembourg (an outlier in several variables) and Malta (the country for which some statistical data were unavailable). It is known that the EU member states adopted all 17 SDGs of the 2030 Agenda for Sustainable Development [1] and have implemented, since 2015, the main strategic measures related to digitalization that aimed to build the EU digital society (e.g., "A Digital Single Market Strategy for Europe) [84]. In addition, in order to highlight the progress made by EU countries regarding digital performance [85], the Digital Economy and Society Index (DESI) has been calculated since 2015. Therefore, the 2018–2019 period allowed the illustration of the entrepreneurship ecosystem in the context of digitalization of an economy and society and its effect on the achievement of the SDGs at the level of EU countries.

The statistical data on the analyzed variables were collected from the Eurostat Database [86], European Innovation Scoreboard Database [87], Global Entrepreneurship and Development Institute Database [88], DESI Datasets [89], Autio et al. [26], World Economic Forum Reports [59,90], Portulans Institute Reports [82,91] and SDSN and IEEP Reports [92,93]. The results of the statistical descriptions of the analyzed variables—mean, maximum and minimum values and standard deviation—are presented in Table 1.

Table 1. Descriptive statistics (n = 25).

Variables	Minimum	Maximum	Mean	Std. Deviation
Digital, innovative and productive entrepreneurship				
SMEs introducing product innovations (SMEs_p_innov) [1]	6.19 (RO)	38.41 (EL)	26.178	9.377
SMEs introducing business process innovations (SMEs_b_innov) [1]	7.73 (RO)	51.97 (EL)	36.888	12.556
Global Entrepreneurship Index (GEI) score (0–100 values)	28.94 (BG)	76.80 (DK)	52.899	13.912
European Index of Digital Entrepreneurship Systems (EIDES) score (0–100 values)	25.30 (BG)	73.95 (DK)	45.080	15.256
Digitalization, competitiveness and economic development				
Digital Economy and Society Index (DESI) score (0–100 values)	35.11 (BG)	70.22 (FI)	51.336	9.878
Network Readiness Index (NRI) score (0–100 values)	54.82 (RO)	82.70 (SE)	67.344	8.773
Global Competitiveness Index 4.0 (GCI 4.0) score (0–100 values)	61.00 (HR)	82.40 (NL)	71.754	6.686
GDP per capita in PPS * (as a percentage of EU-27 average = 100%)	52.00 (BG)	192.00 (IE)	95.640	30.450
Sustainable development goals				
Total SDGs Index [2] score (0–100 values)	56.45 (BG)	80.23 (SE)	68.895	6.807
SDG 8 [3] score (0–100 values)	48.02 (EL)	89.50 (DE)	74.778	10.928
SDG 9 [4] score (0–100 values)	21.92 (BG)	91.66 (SE)	56.742	22.768

Note: [1] Share of total SMEs (%). [2] Total SDGs Index scores = average of all 17 SDGs scores. [3] "Promote sustained, inclusive and sustainable economic growth, full and productive employment and decent work for all". [4] "Build resilient infrastructure, promote inclusive and sustainable industrialization and foster innovation". * PPS = purchasing power standards. BG—Bulgaria; DE—Germany; DK—Denmark; EL—Greece; FI—Finland; HR—Croatia; IE—Ireland; NL—Netherland; RO—Romania; SE—Sweden. Source: own calculations based on [26,59,82,86–93].

3.2. Measures

3.2.1. Digital, Innovative and Productive Entrepreneurship

In order to highlight the *multidimensional nature of entrepreneurship* in the analyzed countries, this study used three complex indicators: the Global Entrepreneurship Index (GEI) for productive entrepreneurship, EIDES for digital entrepreneurship and the innovative SMEs for innovative entrepreneurship.

The quality of entrepreneurship in a country as well as the level of depth of the entrepreneurial ecosystem which fosters it were measured by the Global Entrepreneurship Index (GEI) [52]. Thus, the GEI is a composite index which includes 3 sub-indices (entrepreneurial abilities, attitudes and aspirations) and 14 pillars of the entrepreneurship ecosystem [52], which include both an individual and institutional component that can be connected to the micro- and the macro-level aspects of entrepreneurship. An overall GEI score close to 100% reflects highly productive entrepreneurship and a healthy entrepreneurship ecosystem which is characterized by the productive use of resources and, consequently, an increase in an economy's ability to generate wealth and jobs [94]. According to the statistical data for the 2018–2019 period (Table 1), at the EU-25 level, Denmark was the leader with a GEI maximum value of 75.9%. The lowest value (28.9%) was recorded by Bulgaria, which reflects that this country has an extremely low level of productive entrepreneurship and is very far from a healthy entrepreneurship ecosystem.

For analyzing digital entrepreneurship, we used the *EIDES*, provided by Autio et al. [26]. This composite index assesses the physical and digital prerequisites for stand-up, start-up and scale-up ventures in the EU member states, and it can provide a useful in-depth description of a country's systems of entrepreneurship [22] as an instrument that can help comprehend and evaluate the degree of the digital entrepreneurial ecosystem. The EIDES was developed based on the general and systemic framework conditions at the national level [26]. The values of the overall EIDES score are on a scale from 0 (lowest) to 100 (highest). The average of the EIDES scores for the 2018–2019 period at the EU-25 level was of 45.08 points, which was 28.87 points less than Denmark, the best EIDES performer in the EU. The minimum value was recorded by Bulgaria (Table 1).

Other complex indicators that were used in order to capture entrepreneurship, especially the innovative type, were SMEs introducing product innovations and SMEs introducing business process innovations as a share of the total SMEs. According to the EC Report [95], these indicators form one of the twelve innovation dimensions (the innovators dimension) of the European Innovation Scoreboard, which assesses the innovation performance of the EU member states. Statistical data were obtained from the EIS Database [87] for the 2018–2019 period. It is noticeable that the mean value of the share of SMEs introducing business process innovations (which combine process, marketing and organizational innovations) was higher than the mean value of the share of SMEs introducing product innovations in EU-25 countries (36.88% against 26.17%, respectively; Table 1). As can be seen in Table 1, the mean values of these indicators indicate large differences between the EU-25 countries. Greece was the leader of both types of innovative SMEs (52% and 36%, respectively). Romania recorded the lowest values on both indicators (7.73% and 6.19%, respectively).

3.2.2. Digitalization, Competitiveness and Economic Development

The degree of digitalization of the economies and societies at the EU level was analyzed based on two complex indexes: the DESI and the Network Readiness Index (NRI). The *DESI,* measured by the European Commission [85] since 2015, combines 37 indicators into 5 main interconnected dimensions: connectivity, use of the internet, human capital, integration of digital technology (business digitization and e-commerce) and digital public services. The DESI scores are on a scale from 0 (lowest) to 100 (highest). The statistical data for the 2018–2019 period (Table 1) show that the highest score was recorded by Finland (70.02), while Bulgaria recorded the weakest score (35.1). The *NRI* was initially launched by the World Economic Forum in 2002 in order to evaluate how information and communication technologies (ICTs) influence the development and competitiveness of nations [91]. Starting with 2019, the NRI was redesigned based on a renewed methodology by the Portulans Institute's researchers to reflect how digital technology and people can be used in an efficient governance system so that they influence our society, economy and the environment in a positive way [91], thus accelerating the achievement of the SDGs. Therefore, in our analysis, we used the NRI scores for the 2018–2019 period. The NRI integrates 60 indicators into 4 fundamental dimensions: technology, people, governance and impact (economy, quality of life and SDG contribution) [82]. The NRI scores are based on a 0–100 scale, where 0 is the worst and 100 is the best. The average of the NRI scores for the 2018–2019 period at the EU-25 level was 67.34 points, being 15.36 points less than Sweden, the best NRI performer in the EU. Romania scored the lowest in terms of network readiness (54.82) at the EU level (see Table 1).

The economic development and the level of competitiveness of the national economy were analyzed based on two main indicators: GDP per capita in PPS (as a percentage of the EU-27 average = 100%) and Global Competitiveness Index 4.0 (GCI 4.0). The statistical data (Table 1) show the significant differences between EU countries in terms of the level of GDP per capita as proxies for economic development, which ranged from 52%, in Bulgaria, to 192%, in Ireland (EU-28 = 100%).

The GCI 4.0 was introduced by the World Economic Forum in 2018 and gives a comprehensive map of the driving forces which generate economic growth, high productivity and human development in the Fourth Industrial Revolution era [59]. This complex index is calculated based on 103 individual indicators which are organized into 12 pillars:

infrastructure, institutions, macroeconomic stability, health, skills, labor market, product market, market size, financial system, business dynamism, ICT adoption and innovation capability [59]. A country's performance in terms of competitiveness is reported as a progress score on a scale from 0 to 100. The maximum value (100) is "the frontier", the case in which an issue no longer hampers productivity growth, and consequently, every country should set moving closer to the frontier as a key objective in the overall GCI 4.0 as well as for each of its components. An average GCI 4.0 score of 71.75 for the whole EU-25 countries and for the 2018–2019 period (see Table 1) reflects the EU competitiveness deficit. The Netherlands, the best performer (82.4), still fell 17.6 points short of the frontier. The worst performance in the GCI 4.0 was produced by Croatia (61 points).

3.2.3. Sustainable Development Goals

Achievements in sustainable development at the EU level were emphasized based on three complex indicators (Total SDG Index score (Total SDGs), Sustainable Development Goal 8 Index score (SDG 8) and Sustainable Development Goal 9 Index score (SDG 9)) calculated by the Sustainable Development Solutions Network (SDSN) in cooperation with IEEP [92,93] for the last 2 years (2018–2019). We mention that at EU level, Eurostat has released an SDG dataset and "Sustainable development in the European Union" reports annually since 2016 (e.g., [80]), but this report does not review the overall EU performance based on time-bound targets and does not predict how much "distance to targets" individual EU member states still need to cover to reach the SDGs [96]. Therefore, we used the Total SDGs Index score, SDG 8 score and SDG 9 score in order to highlight each country's performance on a scale from 0 to 100 [92]. These indicators can show the percentages toward achievement of the SDGs (for the 17 goals overall and each goal).

The Total SDG Index score for each EU country is the result of the average of all 17 SDG scores. Every SDG score has an equal weight, underlining the commitment of policymakers to treat all SDGs both equally and indivisibly [93]. The score of SDG 8, "Promote sustained, inclusive and sustainable economic growth, full and productive employment and decent work for all", is grounded in eight indicators which are mainly related to employment, unemployment, disposable income, quality of work and in-work poverty. The score of SDG 9, "Build resilient infrastructure, promote inclusive and sustainable industrialization and foster innovation", is based on eight indicators which mainly cover R&D expenditure and personnel, digital skills, internet connectivity, digital infrastructure, etc. [92].

The average scores for Total SDGs, SDG 8 and SDG 9 for the 2018–2019 period showed that the EU-25 countries achieved 74.78% for SDG 8, 68.9% for Total SDGs and only 56.74% for SDG 9 (see Table 1). As can be seen in Table 1, there are significant differences among the EU-25 countries in terms of all SDG scores analyzed. Sweden tops both the Total SDG Index and SDG 9 with maximum values of 80.23% and 91.66%, respetively. As for SDG 8, Germany was leader in the EU-25 countries, while Greece was the worst performer. The largest discrepancies among EU countries were recorded in terms of SDG 9, with a minimum value of 21.92 (Bulgaria) and a maximum value of 91.66 (Sweden). Thus, Bulgaria is very far from achieving this SDG and also in terms of Total SDGs (minimum value of 56.45%).

3.3. Statistical Methods

We used descriptive statistics, correlation and regression analysis to test the first five research hypotheses (H1–H5). The intensity of the link between the analyzed variables was evaluated based on the Pearson correlation coefficient (r). The calculation of the Pearson correlation coefficient is given in Equation (1):

$$r_{x,y} = \frac{S(x,y)}{Sx \cdot Sy} = \frac{\sum_i^n (xi - \overline{x}) \cdot (yi - \overline{y})}{\sqrt{\sum_i^n (xi - \overline{x})^2} \cdot \sqrt{\sum_i^n (yi - \overline{y})^2}} \quad (1)$$

where S(x,y) is the sample covariance, Sx·Sy is the product of the sample standard deviations of X and Y, respectively, and $\overline{x}$ and $\overline{y}$ are the sample means of X and Y, respectively.

A value of this coefficient closer to −1 or 1 means a stronger negative or positive correlation [97,98].

As opposed to correlation analysis, which indicates only the strength of association between the analyzed variables, regression analysis shows the difference between the dependent and independent variables [99]. Moreover, it reflects the impact of the independent variable on the dependent variable. Therefore, to identify a functional relation between the analyzed variables (DESI, NRI, EIDES, SDG 8, SDG 9, Total SDGs, etc.), we used simple linear regression analysis according to Equation (2):

$$Y = \alpha + \beta \times X + \varepsilon \quad (2)$$

where Y is the dependent variable, X is the explanatory variable, α and β are regression coefficients and ε is the residual or error.

The least squares method was used to estimate the regression coefficients. The validity of the regression model was assessed based on the Fisher–Snedecor (F) statistic. For checking if the errors (residual) of the regression model were affected by autocorrelation, the Durbin–Watson test was applied. A lack of autocorrelation between errors was identified when the values of the Durbin–Watson (DW) test ranged from 1.5 to 2.5. The coefficient of determination (R^2), based on which the quality of the prediction is assessed, indicates to what extend the variance in the dependent variable is explained by the independent variables [99].

Principal component analysis (PCA) and cluster analysis (CA) were employed to test H6. *In the first stage,* considering the main advantage of the PCA being reducing the dimensionality of a dataset which includes a large number of interrelated variables (in our case, 11 variables; see Table 1) to a few principal components or factors that reflect the whole information comprised in the original data [99–101], we used this multivariate technique (PCA with Varimax rotation and Kaiser normalization). The number of principal components was determined based on multiple criteria: Catell's scree plot criterion, the Kaiser criterion or eigenvalue-greater-than-one rule and the percentage of cumulative variance, which retains only those components that indicate a large percentage (between 70 and 90%) of the total variation of the initial variables [99,100].

The choice of using the PCA in this research was based first on the main objective of this method: to reduce the dimensionality of a set of data with many variables while preserving as much statistical information as possible and to create an uncorrelated "new" dataset (a linear combination of variables) for the subsequent multivariate analysis (such as cluster analysis). In the same vein, Kovács et al. [98] compared the PCA "to a shadow game, where the shadow image of a complex spatial shape is projected onto a planar surface so that the characteristic properties of the figure are lost as little as possible" [98] (p. 8). Compared with other data reduction techniques (e.g., factor analysis), PCA is more efficient, as it allowed us to find the components that maximized the variance, which in turn were used in the cluster analysis. Secondly, it was taken into consideration that this multivariate statistical method can contribute to solving the drawbacks generated by the use of different measurement units for the initial variables and the high variations of the covariance coefficients [99,100,102]. Last but not least, PCA and cluster analysis have been widely used by researchers to explore the social and economic differences and similarities between different nations [13,16,103].

In the second stage, the principal components obtained were used in the cluster analysis to identify the relatively homogenous groups of EU countries. The number of the cluster was determined based on hierarchical cluster analysis. Ward's method and the Euclidian distance were employed. Then, the structure of the clusters was based on k-means cluster analysis [102,103]. Thus, these complex statistical methods of data analysis were applied to classify the EU-25 countries and provide a comparative view of the interplay between all 11 variables (see Table 1).

For data processing, IBM SPSS Statistics 26.0 (IBM, Armonk, NY, USA) was used.

4. Results and Discussion

As shown in the previous section, based on descriptive analysis (Table 1), there were large differences between the EU-25 countries in terms of each of the variables analyzed (GEI, EIDES, SMEs_b_innov, SMEs_p_innov, DESI, NRI, GCI 4.0, GDP/capita, Total SDGs, SDG 8 and SDG 9), with the eastern and southern countries generally lagging behind the western and northern countries. Bearing this empirical context in mind, we focused on the statistical analysis of interrelations between (productive, innovative and digital) entrepreneurship and economic development as well as competitiveness and sustainable development in the framework of a digital economy and society in order to find out if these gaps could be explained by the impact of entrepreneurship.

The results of the correlation analysis (Table 2) highlighted that in all EU-25 member states, productive entrepreneurship (measured by the GEI) was strongly positively correlated with both the GDP per capita (r = 0.853) and GCI 4.0 (r = 0.911). The same positive correlation (but of a lower intensity) was set between innovative entrepreneurship (SMEs introducing product innovations and SMEs introducing business process innovations as a share of the total SMEs) on the one hand and, on the other hand, GDP per capita (r = 0.447 and r = 0.464, respectively) and GCI 4.0 (r = 0.449 and r = 0.400, respectively). Additionally, a positive link was identified between innovative entrepreneurship and productive entrepreneurship (r = 0.502 and r = 0.454, respectively), confirming that innovation and entrepreneurship are interlinked [35,53,54]. Therefore, in the EU countries where productive and innovative entrepreneurship was higher (especially the advanced economies of the EU), the level of economic development and competitiveness was also higher, and vice versa. This confirms Hypothesis 1. These findings were sustained by other studies [39,56,57,94], which have highlighted the positive influence of entrepreneurship, characterized by a high level of innovativeness and quality, on macroeconomic performance (expressed by GDP/capita, GCI, total factor productivity, employment, etc.).

Table 2. Correlation matrix.

Pearson Correlation (r)	1	2	3	4	5	6	7	8	9	10	11
1. SMEs_p_innov	1										
2. SMEs_b_innov	0.951 **	1									
3. GEI	0.502 *	0.454 *	1								
4. EIDES	0.500 *	0.429 *	0.940 **	1							
5. DESI	0.449 *	0.378	0.846 **	0.928 **	1						
6. NRI	0.538 **	0.465 *	0.947 **	0.985 **	0.913 **						
7. GCI 4.0	0.449 *	0.400 *	0.911 **	0.945 **	0.827 **	0.969 **	1				
8. GDP/capita	0.447 *	0.464 *	0.853 **	0.779 **	0.687 **	0.770 **	0.765 **	1			
9. Total SDGs	0.391	0.338	0.789 **	0.816 **	0.794 **	0.850 **	0.806 **	0.560 **	1		
10. SDG 8	0.356	0.269	0.767 **	0.802 **	0.784 **	0.805 **	0.731 **	0.572 **	0.733 **	1	
11. SDG 9	0.586 **	0.543 **	0.890 **	0.923 **	0.809 **	0.945 **	0.936 **	0.749 **	0.848 **	0.724 **	1

Note: ** Correlation is significant at the 0.01 level (2-tailed). * Correlation is significant at the 0.05 level (2-tailed). Source: own calculations based on [26,59,82,86–93].

The results from Table 2 and Figure 1 illustrate that productive entrepreneurship (GEI) was very strongly positively associated with the DESI (r = 0.846) and NRI (r = 0.947). SMEs introducing product innovations were moderately linked both with the DESI (r = 0.449) and NRI (r = 0.538). In the case of SMEs introducing business process innovations, a moderately positive and statistically significant link was identified only with the NRI (r = 0.465). Therefore, Hypothesis 2—productive and innovative entrepreneurship is higher in the EU countries with a high degree of digitalization of economy and society—was confirmed and supported by other empirical studies [26,29].

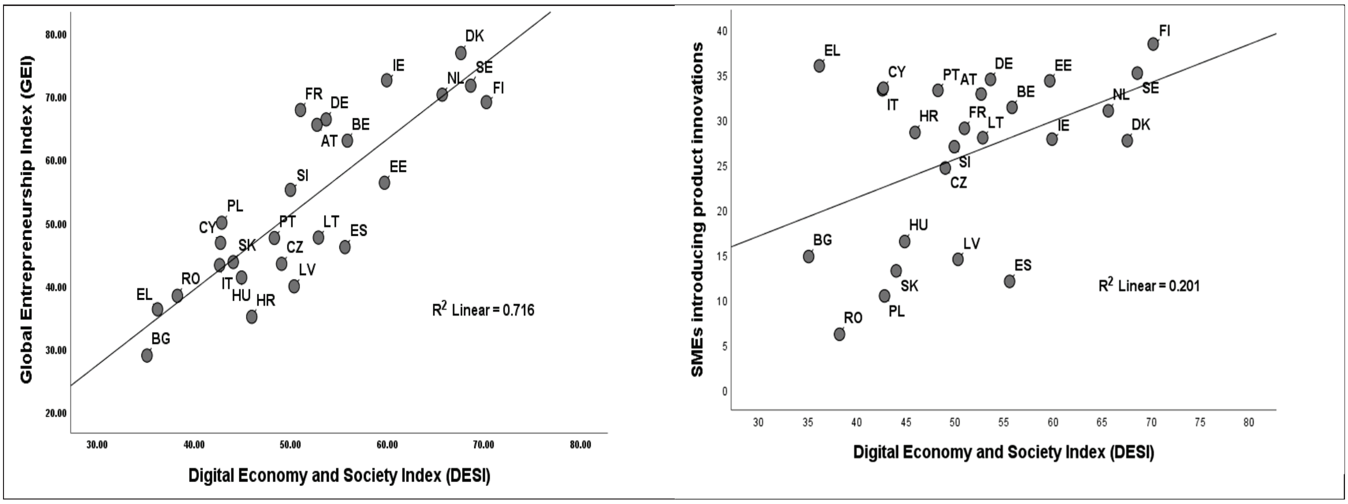

Figure 1. Positive link between productive and innovative entrepreneurship and the digitalization of an economy and society. Source: own calculations based on [59,87,89,90].

Moreover, the GDP per capita and GCI 4.0 had a positive correlation with the DESI as well as with the NRI (Table 2), which means that high economic development and high national competitiveness go hand in hand with highly digitalized countries. Therefore, it is confirmed that the economic performance of countries has been significantly influenced by developments in digital technologies and their adoption [29,64].

Corroborating the positive bivariate link between productive entrepreneurship and innovative entrepreneurship with the positive bivariate link between the EIDES and GEI (r = 0.940) on the one hand and the EIDES and innovative SMEs (r = 0.500 and r = 0.429, respectively) on the other hand proves that, in EU member states, digital entrepreneurship is interrelated with productive and innovative entrepreneurship. Therefore, Hypothesis 3 was supported, and it is in line with [26], which showed a strong positive link between the EIDES and GEI in the context of European economies.

In order to test Hypothesis 4, simple linear regression analysis (Figure 2) was applied. The results pointed out that in the EU countries, digital entrepreneurship (measured by the EIDES) was positively influenced by the DESI (β = 0.928). The regression model was statistically significant (F (1, 23) = 142.803; p = 0.000; Durbin–Watson statistic = 2.148) and accounted for 86.1% of the variance in digital entrepreneurship (R^2 = 0.861, adjusted R^2 = 0.855).

We also conducted an analysis of the impact of the NRI on digital entrepreneurship (Figure 2). The estimated regression model was statistically significant (F (1, 23) = 757.596; p = 0.000; Durbin–Watson statistic = 2.281) and accounted for 97.1% of the variance in digital entrepreneurship (R^2 = 0.971, adjusted R^2 = 0.969). Therefore, it was found that the NRI positively influenced digital entrepreneurship (β = 0.985).

These findings emphasize that in the EU countries where the DESI and NRI are higher, the level of digital entrepreneurship is also high, and consequently, the degree of digitalization of the economy and society was identified as a key driver of digital entrepreneurship, confirming Hypothesis 4. Our results confirmed previous research findings in the context of European economies [20], which highlighted a positive impact of the technology readiness factors (ICT investment and ICT usage by businesses) on technology-driven entrepreneurship. This fact reflects the need to increase the access to and use of digital technologies for both people and business through a more effective governance structure to create an adequate framework for the development of digital entrepreneurship.

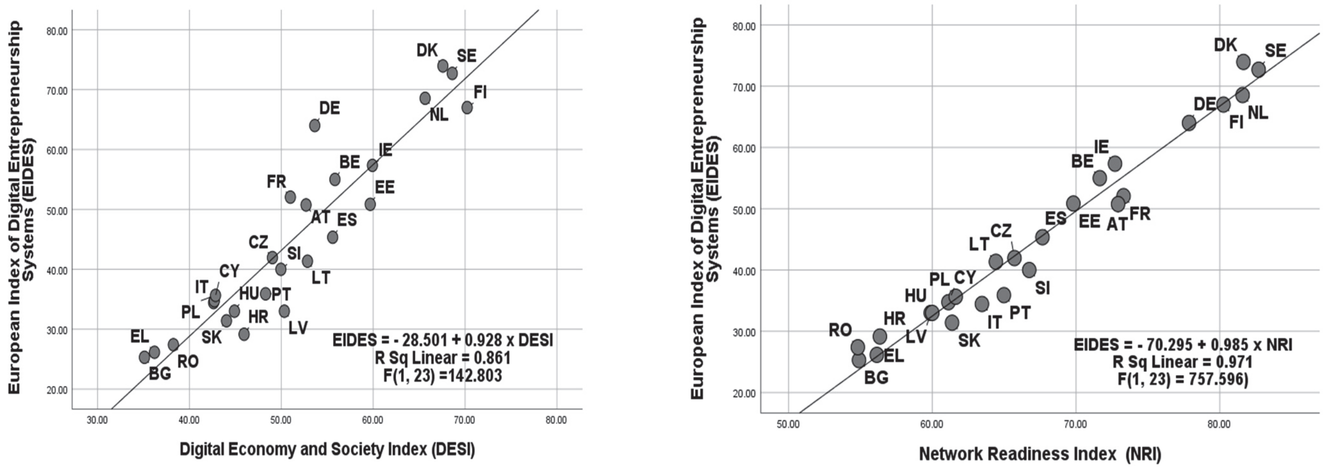

Figure 2. Positive link between digital entrepreneurship and the digitalization of an economy and society. Source: own calculations based on [26,82,89,91].

As noted in Table 2, two of the four variables which reflect the level of entrepreneurship (the GEI and EIDES) were significantly and positively correlated with variables specific to the achievement of SDGs (Total SDGs, SDG 8 and SDG 9). In the case of innovative SMEs, only SMEs_p_innov was significantly positively correlated with SDG 9. In addition, a higher intensity for the EIDES–SDGs link than both the GEI–SDGs link and SMEs_p_innov–SDGs link can be noticed (Table 2).

If we focus on the EIDES–SDGs link (Table 2 and Figure 3), the results of the correlation analysis highlight that the EIDES was very strongly correlated with SDG 9 (r = 0.923), Total SDGs (r = 0.816) and SDG 8 (r = 0.802). Furthermore, the regression models (models 1–3, Table 3) pointed out a significantly positive impact by digital entrepreneurship on sustainability (Total SDGs, SDG 8 and SDG 9). All models were statistically significant (F (1, 23) = 45.787, p = 0.000, R^2 = 0.666 for model 1; F (1, 23) = 41.434, p = 0.000, R^2 = 0.643 for model 2; F (1, 23) = 131.639, p = 0.000, R^2 = 0.851; see Table 3). The EIDES accounted for 66.6% of the variance in Total SDGs, 64.3% of the variance for SDG 8, and 85.1% of the variance for SDG 9. Thus, the high level of sustainable development in the EU member states can be explained by the existence of a high level of digital entrepreneurship. Based on these results, Hypothesis H5—digital entrepreneurship has a positive impact on the achievement of SDGs (Total SDGs, SDG 8 and SDG 9)—is confirmed.

Table 3. Regression results: the impact of digital entrepreneurship (EIDES) on the achievement of SDGs (Total SDGs, SDG 8 and SDG 9).

Models		Unstandardized Coefficients		Standardized Coefficients	t-Statistics	Sig.
		B	Std. Error	Beta		
Model 1 [1] EIDES–Total SDGs	Constant	52.485	2.555		20.542	0.000
	EIDES	0.364	0.054	0.816	6.767	0.000
Model 2 [2] EIDES–SDG 8	Constant	48.884	4.238		11.534	0.000
	EIDES	0.574	0.089	0.802	6.437	0.000
Model 3 [3] EIDES–SDG 9	Constant	−5.331	5.700		−0.935	0.359
	EIDES	1.377	0.120	0.923	11.473	0.000

[1] Dependent variable: Total SDGs; R^2 = 0.666, adjusted R^2 = 0.651; std. error of the estimate = 4.020787; Durbin–Watson statistic = 2.266; F (1, 23) = 45.787, $p < 0.001$. [2] Dependent variable: SDG 8; R^2 = 0.643, adjusted R^2 = 0.628; std. error of the estimate = 6.66952; Durbin–Watson statistic = 1.637; F (1, 23) = 41.434, $p < 0.001$. [3] Dependent variable: SDG 9; R^2 = 0.851, adjusted R^2 = 0.845; std. error of the estimate = 8.96971; Durbin–Watson statistic = 2.236; F (1, 23) = 131.639, $p < 0.001$. Source: own calculations based on [26,92,93].

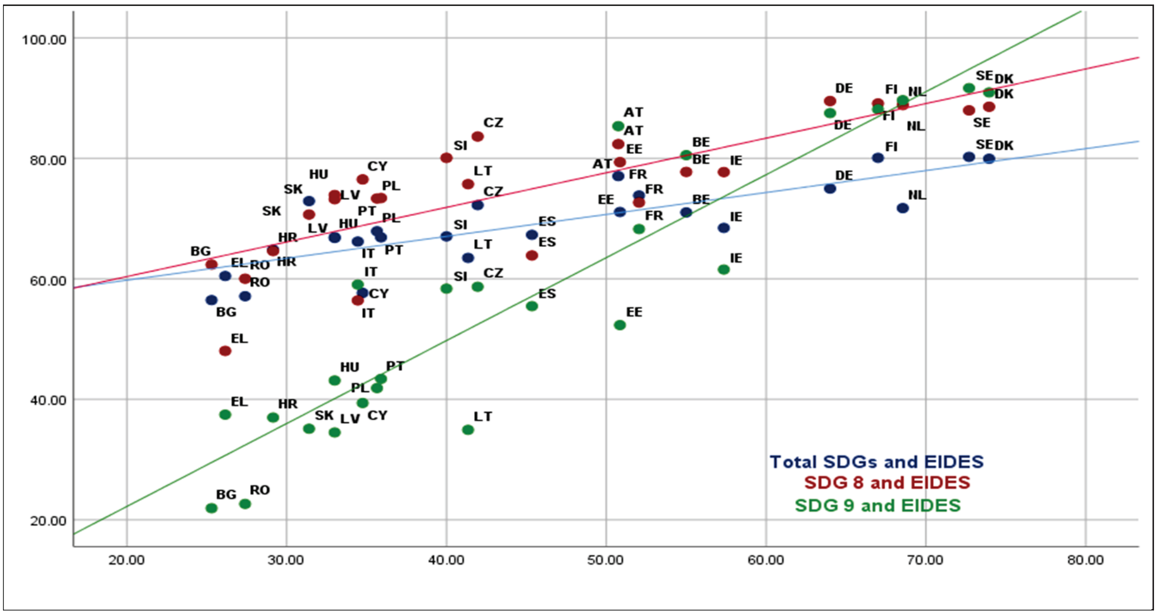

Figure 3. Positive link between digital entrepreneurship (EIDES) and sustainable development goals (Total SDGs, SDG 8 and SDG 9). Source: own calculations based on [26,92,93].

In order to find out if there were differences and common features between the EU countries based on their interrelations between (digital, innovative and productive) entrepreneurship, digitalization, economic development, national competitiveness and the Sustainable Development Goals (Hypothesis 6), we took into account the cumulative influence of all 11 variables (GEI, EIDES, SMEs_b_innov, SMEs_p_innov, DESI, NRI, GCI 4.0, GDP/capita, Total SDGs, SDG 8 and SDG 9; see Tables 1 and 2), and we employed complex statistical methods of data analysis (PCA and CA). As shown above, statistically significant correlations were identified between the initial variables included in the PCA (see correlation matrix, Table 2).

The overall Measure of Sampling Adequacy (MSA) was checked based on the Kaiser–Meyer–Olkin (KMO) indicator, whose value of 0.828 exceeded the minimum requirement of 0.50 [102]. Bartlett's test of sphericity (approximate. chi-square = 423.574, p-value = 0.000) supported the overall MSA results, a fact which indicates the suitability of the variables set for the principal component analysis.

Based on the PCA with the Varimax rotation method with Kaiser normalization (rotation converged in 3 iterations), the 11 initial variables were divided into two components which explained 87.615% of the total variance (Table 4).

Table 4. Total variance and eigenvalues explained.

Component	Initial Eigenvalues			Extraction Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	8.202	74.567	74.567	8.202	74.567	74.567
2	1.435	13.048	87.615	1.435	13.048	87.615
...	...	...	...			
11	0.003	0.028	100.000			

Note: The extraction method was principal component analysis.

The first principal component (PC1) was strongly and positively correlated with 9 of the original variables and accounted for 74.567% of the total variance in the observed variables (Tables 4 and 5). Thus, all nine variables had a positive contribution in the formation of this principal component and reflected the level of digital and productive entrepreneurship (EIDES and GEI), degree of digitalization (DESI and NRI), national competitiveness (GCI

4.0), economic development (GDP/capita) and Sustainable Development Goals (Total SDGs, SDG 8 and SDG 9). *The second principal component* (PC2), which explained 13.048% of the total variance (Table 4), was strongly positively correlated with two variables (Table 5), reflecting the level of innovative entrepreneurship (SMEs_p_innov and SMEs_b_innov).

Table 5. Principal components for EU-25 countries (rotated component matrix).

Initial Variables	PC1	PC2
European Index of Digital Entrepreneurship Systems (EIDES)	0.948	0.258
Network Readiness Index (NRI)	0.947	0.294
Global Competitiveness Index 4.0 (GCI 4.0)	0.927	0.230
Global Entrepreneurship Index (GEI)	0.914	0.290
Digital Economy and Society Index (DESI)	0.898	0.202
SDG 9	0.875	0.386
Total SDGs	0.863	0.153
SDG 8	0.846	0.087
GDP/capita	0.743	0.337
SMEs introducing business process innovations (SMEs_b_innov)	0.197	0.969
SMEs introducing product innovations (SMEs_p_innov)	0.267	0.943

Note: The extraction method was PCA, and the rotation method was Varimax with Kaiser normalization. The rotation converged in 3 iterations.

In the next step, the two principal components (PC1 and PC2) were used in the cluster analysis to classify the EU-25 member states. The 3 formed clusters were statistically significant according to the results of the ANOVA analysis (for PC1, F (3, 21) = 23.319, $p < 0.001$; for PC2, F (3, 21) = 66.709, $p < 0.001$; Figure 4 and Table 6).

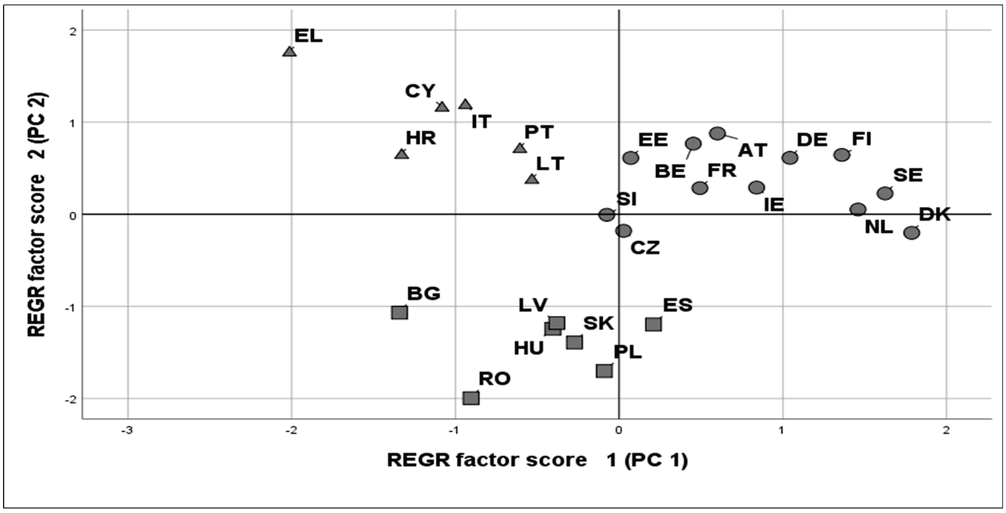

Figure 4. EU clusters based on PCA and CA.

Table 6. The results of the cluster analysis: final cluster centers and ANOVA.

	Final Cluster Centers			ANOVA					
				Cluster		Error			
	Cluster 1	Cluster 2	Cluster 3	Mean Square	df	Mean Square	df	F	Sig.
PC 1	0.807	−0.455	−1.084	8.154	2	0.350	22	23.319	0.000
PC 2	0.331	−1.397	0.968	10.301	2	0.154	22	66.709	0.000

Cluster 1 was powerfully correlated with PC1 (0.807, Table 6) and included 12 countries (Denmark, Finland, Ireland, the Netherlands, Sweden, Austria, Belgium, Germany, France, the Czech Republic, Estonia and Slovenia). All countries (except the Czech Republic, Estonia and Slovenia) are "old EU member states", being characterized by a high level of digitalization of the economy and society (DESI and NRI) and of national competitiveness (GCI 4.0). Compared with the other clusters, this cluster's scores for digital entrepreneurship (EIDES) and productive entrepreneurship (GEI) were the highest. Additionally, the economic development (GDP/capita) and sustainable development goals (Total SDGs, SDG 8 and SDG 9) scored the highest (Figures 5–7). These results emphasize that a favorable (digital and non-digital) infrastructure and institutions specific to the countries in this cluster (particularly in the northern countries of the EU) stimulated both productive and digital entrepreneurship, which in turn generated inclusive and green economic growth and development. As displayed in Figure 7, SDG 8 and SDG 9 recorded higher values related to the overall SDGs (Total SDGs), a fact which shows that the countries in this cluster performed better in terms of socioeconomic goals, but they have to address major challenges in achieving other SDGs linked to climate action, responsible consumption and production and biodiversity. In spite of these achievements, our results show that this cluster displayed an EIDES average value per cluster of 57.8% (of 100%, "the frontier") and an attainment value of around 74% of the Total SDGs established for 2030. These findings suggest that there is room for improvement both in terms of digital entrepreneurship and the SDGs. To achieve all SDGs in these countries, digital technologies should be adopted, integrated and exploited by business and society so that they ensure a dynamic balance between the economic, social and ecological dimensions of sustainability.

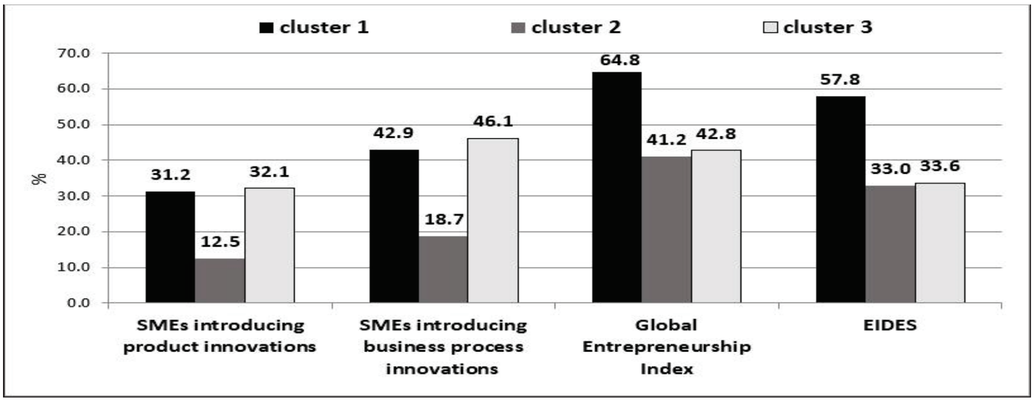

Figure 5. Digital, innovative and productive entrepreneurship (mean values per cluster). Source: own calculations based on [26,87,88].

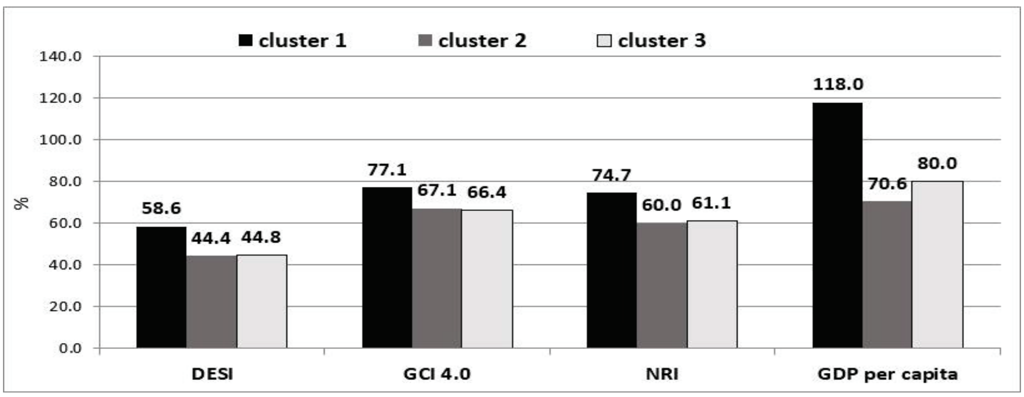

Figure 6. Digitalization, economic development and competitiveness (mean values per cluster). Source: own calculations based on [59,82,86,89–91].

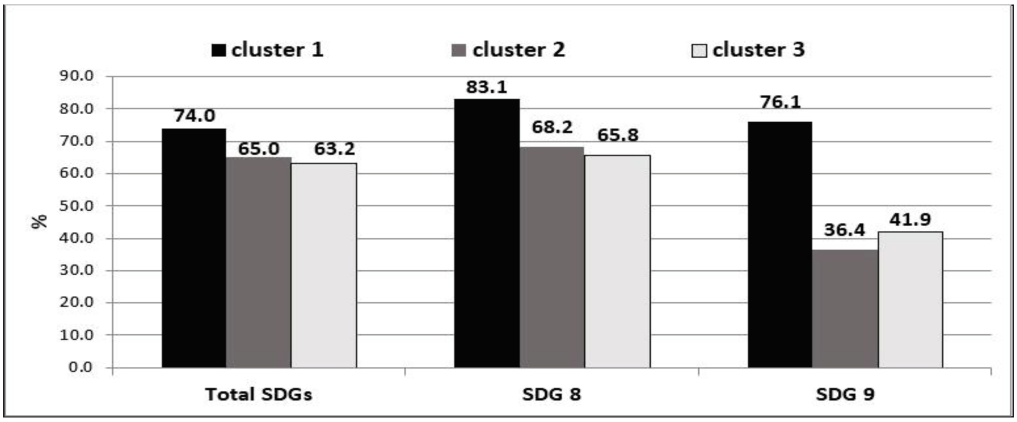

Figure 7. Sustainable Development Goals (SDGs) (mean values per cluster). Source: own calculations based on [92,93].

Moreover, the countries in this cluster showed high heterogeneity. Inclusion of the Czech Republic, Estonia and Slovenia in this cluster was due to higher values for indicators specific to the digitalization of the economy (DESI and NRI), digital entrepreneurship (EIDES) and sustainable development compared with countries in other clusters, but they were smaller compared with the peer countries in the cluster.

Cluster 2 was very strongly and negatively correlated with PC 2 (−1.397, Table 6) and weakly and negatively correlated with PC 1 (−0.455, Table 6), thus being placed in the third quadrant (except Spain). This cluster contained seven countries (Bulgaria, Poland, Hungary, Romania, Slovakia, Latvia and Spain) mainly characterized by their lowest values in terms of innovativeness in business (SMEs introducing product innovations as well as SMEs introducing business process innovations) and SDG 9 (Figures 5–7).

Spain's position in the fourth quadrant (Figure 4) was due to the best value of the EIDES, DESI, NRI, GCI 4.0, GDP/capita, and SDG 9 (all variables defining PC1) compared with other countries in the same cluster. Excluding Spain, this cluster included former centrally planned economies and new EU member states, which are still largely lagging behind the EU's mature economy markets. These countries had the lowest level of GDP/capita, which means that they had fewer resources and would be less able to better invest in innovation and in the kinds of institutions and infrastructure that are included in the EIDES and GEI.

Bulgaria and Romania were the countries with the worst values of all 11 variables, both related to peer countries in the cluster and to all EU-25 countries. These large gaps between EU countries and the fragmentation itself can be seen as an obstacle to the adoption and use of EU digital solutions on a large scale. The lowest score for SDG 9 (of 36.4 points) reflects that this cluster obtained 47.83% SDG 9 performance in cluster 1 (best performer) (Figure 7). The poorest performance of this cluster in terms of SDG 9 (Industry, Innovation and Infrastructure), and the existence of very large gaps related to other clusters underlines a necessity to enhance productivity and innovation to boost convergence across EU countries, as also emphasized by the EU report [92].

Cluster 3 included six countries (Croatia, Cyprus, Greece, Italy, Lithuania and Portugal). This cluster was strongly negatively correlated with PC1 (−1.084, Table 6) and strongly positively correlated with PC2 (0.968, Table 6). Thus, some of the indicators which defined PC1 such as the GCI 4.0, Total SDGs and SDG 8 had the lowest values. Among the three clusters, this ranked first in terms of all indicators which defined PC2 (cluster 2 ranked last), and therefore, it was characterized by the highest level of innovativeness of SMEs (Figures 5–7). In spite of this, the countries in this cluster performed worse compared with cluster 1 in terms of national competitiveness, economic development and achievement of Sustainable Development Goals due to resource constraints and a limited capacity to benefit from the economies of scale of SMEs. Between clusters 3 and 2, there were insignificant differences (below 1 point) regarding the indicators which reflect digital entrepreneurship (EIDES), the degree of digitalization of an economy and society (DESI) and national competitiveness (GCI 4.0).

As for the achievement of the overall SDGs (Total SDGs) and SDG 8 ("Decent Work and Economic Growth"), it was noticed that cluster 3 ranked last. The most recent statistical data [86] show that the countries in cluster 3, especially Greece and Italy, had high in-work at-risk-of-poverty rates (15.1% and 15.6%, respectively), high shares of people at risk of poverty or social exclusion in the total population (29% and 21.6%, respectively) and high unemployment rates (17.3% and 10%, respectively), facts which explain the lowest performance of cluster 3 for SDG 8. In these countries, innovative and digital entrepreneurship should be boosted to generate better jobs and higher labor incomes. Consequently, workers would escape poverty, improve their quality of life and respond to ecological challenges through environmentally responsible consumption and production.

The findings above prove that there are differences and common features between EU countries based on their interrelation between (digital, innovative and productive) entrepreneurship, digitalization, economic development, national competitiveness and SDGs. Thus, Hypothesis H6 is confirmed. Therefore, different and specific steps are required to support both a high level of sustainable development and convergence across EU countries.

In order to identify some potential policies for increasing the positive impact of digital entrepreneurship on sustainable development at the EU country level, especially in the case of countries included in cluster 2, we analyzed the pillars based on which the EIDES rests [26]. To obtain a clearer picture of the weakest pillars of the EIDES for the analyzed countries (the pillars of general and systemic framework conditions at the national level for both the digital and non-digital versions), these are centralized in Table 7 based on data from the last EIDES report [26]. According to the data from this report on the EIDES [26], all countries in cluster 2 (except Spain) were included in the "laggards" group. According to the data from Table 7, these countries' EIDES efficiencies relative to the EU leader (Denmark) ranged from 34.35% (Bulgaria) to 48.4% (Poland), which emphasizes the need for policies in order to improve the EIDES score and its impact on the SDGs.

Table 7. The weakest pillars of EIDES.

Countries	EIDES SCORE/ "Efficiency" *	General Framework Conditions		Systemic Framework Conditions	
		Non-Digital Score	Digital Score	Non-Digital Score	Digital Score
Bulgaria	26.9/ 34.35%	Physical infrastructure	Market conditions	Knowledge creation and dissemination	Human capital
Romania	29.5/ 37.67%	Physical infrastructure	Market conditions	Finance	Human capital
Slovakia	33.1/ 42.27%	Formal institutions, regulation, taxation	Formal institution, regulation, taxation	Human capital	Networking and support
Hungary	34.3/ 43.81%	Culture, informal institutions	Culture, informal institutions	Networking and support	Finance
Latvia	34.3/ 43.81%	Physical infrastructure	Market conditions	Knowledge creation and dissemination	Knowledge creation and dissemination
Poland	37.9/ 48.4%	Formal institution, regulation, taxation	Formal institution, regulation, taxation	Knowledge creation and dissemination	Human capital

Note: * Efficiency of EIDES relative to the EU leader (Denmark's EIDES = 78.3) Source: based on [26].

Bulgaria and Romania need to improve their EIDES scores, with both countries displaying EIDES efficiencies relative to the EU leader (Denmark) below 40% (Table 7). The weakest pillars of the EIDES which need to be improved are digital market conditions and non-digital physical infrastructure, as well as digital human capital. As for the latter, statistical data for 2019 (Table 8) proves that Bulgaria and Romania had the lowest levels in the EU in terms of individuals with at least basic digital skills (29.4% and 31% of total individuals, respectively) compared with the EU average of 56.1% and the best EU performer (the Netherlands: 79.4%). These data show that all countries in cluster 2 (especially Romania and Bulgaria) as well as the whole of the EU are very far from meeting the target of 70% in 2025 (according to The European Skills Agenda). Digital skills are recognized as a critical value for working, social interaction and learning on the one side and, on the other side, a key factor for the growth of digital businesses [7,26,80]. Given the aspects above, it is necessary to fill the digital skills gap so that both digital entrepreneurship and sustainable development increase in the analyzed countries.

Table 8. Barriers to EIDES.

Countries	Digital Skills [1]	GERD *		Enterprises with E-Commerce Sales [3]	Ease of Doing Business Score [4] and Rank [5]
		Business Enterprise Sector [2]	Government Sector [2]		
Bulgaria	29	0.56	0.21	11	72.0 (61)
Romania	31	0.28	0.15	12	73.3 (55)
Slovakia	54	0.45	0.17	15	75.6 (45)
Hungary	49	1.11	0.15	15	73.4 (52)
Latvia	43	0.17	0.12	14	80.3 (19)
Poland	44	0.83	0.02	16	76.4 (40)
EU	56	1.46	0.25	20	
Denmark	70	1.82	0.08	34	85.3 (4)

Note: * GERD = gross domestic expenditure on R&D. [1] Individuals who have basic or above basic overall digital skills (% of individuals). [2] Percent of gross domestic product (GDP). [3] Percent of all enterprises without the financial sector (10 persons employed or more). [4] “An economy’s ‘ease of doing business’ score is reflected on a scale from 0 to 100, where 0 represents the lowest and 100 represents the best performance”. [5] Economy’s position relative to that of other economies (1–191 countries). Source: based on [86,104].

The weakest digital market conditions pillars in Romania, Bulgaria and Latvia can be explained by low and very low shares of enterprises with e-commerce sales (compared with the EU-25 average and EIDES performer) (Table 8). Moreover, according to the DESI report [85], in these countries, there is a low and very low level for the Digital Intensity Index, which assesses the degree of using different digital technologies (0–12 digital technologies) at the enterprise level, highlighting that over 55% of businesses invested very little in digital technologies.

As exhibited in Table 7, in Romania, the weakest non-digital systemic framework condition was “finance”, which refers to financing SMEs and domestic credit to the private sector. Therefore, Romania must act in order to improve its access to finance for all entrepreneurs. In the case of Bulgaria, the “knowledge creation and dissemination” non-digital pillar was the weakest, highlighting the need to create a favorable framework, which should assure entrepreneurs’ access to the essential knowledge that drives their business ventures [26]. Moreover, it is necessary to boost the absorptive capacity of business firms to integrate knowledge inputs into new products and services. The “knowledge creation and dissemination” non-digital pillar was also the weakest in the cases of Poland and Latvia (Table 7). One of the main drawbacks of this pillar is the very low level of investment in R&D (GERD) in both the private and public sectors, as shown in Table 8. This result is congruent with other research [9,50,54], which proved that the level of R&D expenditure is often associated with innovation and entrepreneurial performance.

The weakest of the EIDES results within the general framework conditions recorded by Slovakia and Poland were digital and non-digital “formal institutions, regulation, taxation”. Therefore, these countries need to improve the efficiency and quality of their formal institutions and regulations. As seen in Table 8, the statistical data for the ease of doing business (EDB) score and rank point out a large gap between the analyzed countries from cluster 2 and Denmark (the best EIDES performer). Thus, a critical way to improve the institutional framework to harness the productive and digital entrepreneurship and its impact on sustainable development is through the widespread use of electronic systems (electronic tax filing platforms, online business incorporation processes and online procedures related to property transfers), which is, according to the Doing Business Report [104], a common feature of economies that score highest for EDB.

The weakest pillar for Hungary was "culture, informal institutions", which highlights the need to increase the positive cultural and social norms and practices to enhance productive entrepreneurship by making careers in entrepreneurship more attractive and promoting risk taking for business growth [26]. Cultural values, social norms and practices as well as other informal institutions have rapidly been shaped and changed by digitalization, but this impact depends on the capacity of individuals and businesses to harness a country's digital infrastructure. "Networking and support" were the weakest pillars of digital and non-digital framework conditions in Slovakia and Hungary, and therefore these countries need to boost a positive and supportive attitude toward entrepreneurs to national and international networks as well as formal and informal access to resources though social networks, including virtual social networks.

For everyone to benefit from digital technologies and to reduce the gaps between countries, ICT and digital investments are required, but "analog complements" [64] also have to be improved by creating regulations and institutions more favorable to business and entrepreneurship, as well as by improving workers' skills, especially digital skills.

These results are in line with other research papers [37,49,51,105] that highlighted that there is a high heterogeneity of environmental conditions of entrepreneurship and proved the absence of an ideal context. Furthermore, the critical impact of these conditions on the type of entrepreneurship (including digital entrepreneurship) and its effect on the economy and society are stressed. Therefore, it is imperative to turn the barriers to (digital or non-digital) entrepreneurship into strong triggers and driving forces.

5. Conclusions and Main Implications

The fundamental research question of this paper was as follows: can digital entrepreneurship contribute to achievement of the SDGs in the context of the EU digital economy and to what extent? Therefore, this study has shed light on the interrelationship between digital entrepreneurship and productive and innovative entrepreneurship and its impact on achievement of the SDGs in the context of the digital economy in the EU member states for the 2018–2019 period. Moreover, we consider that it is essential to assess the level of digital entrepreneurship in the EU countries and its impact on sustainable development as well as identify both key opportunities and barriers in order to generate more and better digital entrepreneurship as a key driver for sustainable development.

The results of the correlation analysis show that in the EU-25 countries in the 2018–2019 period, a positive link was identified between the productive and innovative entrepreneurship (GEI and innovative SMEs) and macroeconomic performance (GDP/capita and GCI 4.0) on the one hand and between the productive and innovative entrepreneurship and the degree of digitalization of the economy and society (DESI and NRI) on the other hand. Thus, the EU countries (especially the more developed countries) with higher productive and innovative entrepreneurship were characterized by higher economic development and national competitiveness and were also highly digitalized. These findings clearly show that countries with better economic, competitive and digital backgrounds are more likely to be innovative, productive and digital entrepreneurial countries. Through corroborating these results with the strong and positive correlation identified between the GDP/capita, GCI 4.0, DESI and NRI, we conclude that a high level of economic development and national competitiveness goes hand in hand with a high level of digitalization of the economy and society and in turn goes hand in hand with productive and innovative entrepreneurship.

Moreover, our findings prove that at the level of the EU-25 countries, digital entrepreneurship (EIDES) was strongly positively interlinked with productive and innovative entrepreneurship as well as with the degree of digitalization. Therefore, it is argued that digital entrepreneurship is the result of the intersection of digital technologies and entrepreneurship [23,66] and fosters productive and innovative entrepreneurial activities. Our analysis showed that low digital and productive entrepreneurship are predominantly located in the less-digitalized countries, mainly among new EU member states. This result highlights the need to improve the digital infrastructure and the access to and use of digital technologies for both entrepreneurs and businesses in these countries.

Furthermore, the regression analysis results highlighted that the achievement of the Sustainable Development Goals (Total SDGs, SDG 8 and SDG 9) was positively influenced by the level of digital entrepreneurship. Thus, the gaps in the achievement of SDGs between the EU countries can be explained by the level and characteristics of digital entrepreneurship, a fact which highlights the need to implement specific policies to boost digital, innovative and productive entrepreneurship.

Based on the findings of the PCA and cluster analysis, the EU-25 countries were classified into three clusters which confirmed the differences and common features between these countries in terms of the interrelation between (digital, innovative and productive) entrepreneurship, digitalization, economic development, national competitiveness and achievement of the SDGs. Cluster 1, which mainly included the EU-15 countries (except the Czech Republic, Estonia and Slovenia), was the best performer in terms of all 11 analyzed variables (except the innovative SMEs). It is worth mentioning that even if the countries in cluster 1 are at an advantage when it comes to benefiting from the next stage of digital transformation, proactive actions are needed on multiple levels to foster the development and implementation of high-potential technologies, keeping in mind that such a transition requires caution and awareness of the potential risks [106].

The countries included in cluster 3 (Croatia, Cyprus, Greece, Italy, Lithuania and Portugal) were mainly characterized by a high level of innovativeness in their SMEs but were very far from achieving SDG 8 (65.8% of SDG 8 established for 2030) and the Total SDGs (63.2% of the Total SDGs established for 2030). Thus, in these countries, more efficient integration of SMEs into the value chain and market is required. There is a growing need to increase the resources and capacities of the SMEs so they benefit from economies of scale and become more productive. To achieve all SDGs by 2030, especially SDG 8 ("Promote sustained, inclusive and sustainable economic growth, full and productive employment and decent work for all"), as this cluster was the worst performer in terms of the attainment of SDG 8, these countries need to strengthen the link between (digital and non-digital) entrepreneurship, productive activities and decent job creation. Moreover, they also need to respond to ecological challenges.

The poorest performance was obtained by cluster 2 (Bulgaria, Poland, Hungary, Romania, Slovakia, Latvia and Spain), characterized by the lowest level of innovativeness of business, productive entrepreneurship, digitalization of the economy and society with negative consequences on economic development, national competitiveness and achievement in the SDGs. Based on the analysis of the weakest pillars of digital entrepreneurship (EIDES), some specific measures were identified in order to improve digital entrepreneurship and its impact on achievement of the SDGs. Thus, the countries in cluster 2—but not only these countries—need to invest more in R&D and digital technologies to support access to finance for all entrepreneurs and to improve the efficiency and quality of formal institutions and regulations. It is noteworthy that whether and to what extent the EU countries meet the SDGs by 2030 critically depends on how adaptive countries are to future technological changes [107]. Moreover, an improvement in digital skills is a prerequisite, as (digital) human capital lies at the heart of the EU's twin (green and digital) transitions.

Given the results achieved in this work and the significant potential that digital entrepreneurship can have in the attainment of sustainable development, we consider that EU policies should encourage entrepreneurs to adopt, integrate and exploit digital technologies in their businesses so that these become economically viable, socially responsible and environmentally friendly.

Our study contributes to both the entrepreneurship and sustainability literature by providing an empirical approach, which demonstrates not only the positive impact of digital entrepreneurship on the achievement of the SDGs (Total SDGs, SDG 8 and SDG9) but also confirms the assumption that the latter are inextricably intertwined. Furthermore, the findings of this study can be useful for decision makers to formulate policies that stimulate digital entrepreneurship to reach the SDGs.

Limitations and Future Research

First, our analysis was restricted to the period before COVID-19 because of time lags in data creation and release. Moreover, because of the unavailability of statistical data for all 11 variables used in the principal component analysis and cluster analysis, our study was limited to the 2018–2019 period. Therefore, further research should investigate to what extent the COVID-19 pandemic affected the level and dynamic of digital entrepreneurship and its impact on the achievement of the SDGs by 2030. Furthermore, our analysis used only one variable for digital entrepreneurship provided by secondary datasets [26], and therefore future research should focus on taking into consideration more variables on the one side and on directions to build primary information for digital entrepreneurship on the other side. Secondly, in order to assess achievement of the SDGs at the EU level, our analysis was limited to only three indices: Total SDGs, SDG 8 and SDG 9. Therefore, this research paper could be extended in future research by taking into account many other indices related to education (SDG 4), health (SDG 3), climate change (SDG 13) and reduction of inequalities (SDG 10), among others.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. United Nations (UN) General Assembly. *Transforming Our World: The 2030 Agenda for Sustainable Development*; Resolution Adopted by the General Assembly on 25 September 2015; A/RES/70/1; United Nations: New York, NY, USA, 2015. Available online: https://www.un.org/ga/search/view_doc.asp?symbol=A/RES/70/1&Lang=E (accessed on 24 May 2020).
2. Castro, G.D.R.; Fernández, M.C.G.; Colsa, Á.U. Unleashing the convergence amid digitalization and sustainability towards pursuing the Sustainable Development Goals (SDGs): A holistic review. *J. Clean. Prod.* **2021**, *280*, 122204. [CrossRef]
3. Brundtland, G. *Our Common Future: World Commission on Environment and Development (WCED)*; Oxford University Press: Oxford, UK, 1987.
4. European Commission (EC). EU Holistic Approach to Sustainable Development. The EU Approach towards Implementing the UN's 2030 Agenda for Sustainable Development Together with Its Member States. 2021. Available online: https://ec.europa.eu/info/strategy/international-strategies/sustainable-development-goals/eu-holistic-approach-sustainable-development_en (accessed on 4 June 2021).
5. UN General Assembly. *Entrepreneurship for Sustainable Development*; Resolution Adopted by the General Assembly on 21 December 2020; United Nations: New York, NY, USA, 2020. Available online: https://unctad.org/system/files/official-document/ares75d211_en.pdf (accessed on 4 March 2021).
6. Moya-Clemente, I.; Ribes-Giner, G.; Pantoja-Diaz, O. Identifying environmental and economic development factors in sustainable entrepreneurship over time by partial least squares (PLS). *PLoS ONE* **2020**, *15*, e0238462. [CrossRef] [PubMed]
7. Beliaeva, T.; Ferasso, M.; Kraus, S.; Damke, E.J. Dynamics of digital entrepreneurship and the innovation ecosystem: A multilevel perspective. *Int. J. Entrep. Behav. Res.* **2020**, *26*, 266–284. [CrossRef]
8. Cervelló-Royo, R.; Moya-Clemente, I.; Perelló-Marín, M.R.; Ribes-Giner, G. Sustainable development, economic and financial factors, that influence the opportunity-driven entrepreneurship. An fsQCA approach. *J. Bus. Res.* **2020**, *115*, 393–402. [CrossRef]
9. Pradhan, R.P.; Arvin, M.B.; Nair, M.; Bennett, S.E. The dynamics among entrepreneurship, innovation, and economic growth in the Eurozone countries. *J. Policy Modeling* **2020**, *42*, 1106–1122. [CrossRef]
10. Neumann, T. The impact of entrepreneurship on economic, social and environmental welfare and its determinants: A systematic review. *Manag. Rev. Q.* **2021**, *71*, 553–584. [CrossRef]
11. Apostolopoulos, N.; Al-Dajani, H.; Holt, D.; Jones, P.; Newbery, R. Entrepreneurship and the Sustainable Development Goals. In *Entrepreneurship and the Sustainable Development Goals*; Emerald Publishing Limited, Howard House: Bingley, UK, 2018; pp. 1–7.
12. Global Entrepreneurship Research Association (GERA). Global Entrepreneurship Monitor (GEM) 2019–2020 Global Report. Available online: http://www.gemconsortium.org/report (accessed on 20 March 2021).

13. Szabo, Z.K.; Herman, E. Productive Entrepreneurship in the EU and Its Barriers in Transition Economies: A Cluster Analysis. *Acta Polytech. Hung.* **2014**, *11*, 73–94.
14. Du, K.; O'Connor, A. Entrepreneurship and advancing national level economic efficiency. *Small Bus. Econ.* **2018**, *50*, 91–111. [CrossRef]
15. Szerb, L.; Lafuente, E.; Horváth, K.; Páger, B. The relevance of quantity and quality entrepreneurship for regional performance: The moderating role of the entrepreneurial ecosystem. *Reg. Stud.* **2018**, *53*, 1308–1320. [CrossRef]
16. Filser, M.; Kraus, S.; Roig-Tierno, N.; Kailer, N.; Fischer, U. Entrepreneurship as catalyst for sustainable development: Opening the black box. *Sustainability* **2019**, *11*, 4503. [CrossRef]
17. Johnson, M.P.; Schaltegger, S. Entrepreneurship for sustainable development: A review and multilevel causal mechanism framework. *Entrep. Theory Pract.* **2020**, *44*, 1141–1173. [CrossRef]
18. Dhahri, S.; Omri, A. Entrepreneurship contribution to the three pillars of sustainable development: What does the evidence really say? *World Dev.* **2018**, *106*, 64–77. [CrossRef]
19. Hall, J.; Daneke, G.A.; Lenox, M.J. Sustainable development and entrepreneurship: Past contributions and future directions. *J. Bus. Ventur.* **2010**, *25*, 439–448. [CrossRef]
20. Jafari-Sadeghi, V.; Garcia-Perez, A.; Candelo, E.; Couturier, J. Exploring the impact of digital transformation on technology entrepreneurship and technological market expansion: The role of technology readiness, exploration and exploitation. *J. Bus. Res.* **2021**, *124*, 100–111. [CrossRef]
21. Autio, E.; Nambisan, S.; Thomas, L.D.; Wright, M. Digital affordances, spatial affordances, and the genesis of entrepreneurial ecosystems. *Strateg. Entrep. J.* **2018**, *12*, 72–95. [CrossRef]
22. Autio, E.; Szerb, L.; Komlósi, É.; Tiszberger, M. *The European Index of Digital Entrepreneurship Systems*; Nepelski, D., Rossetti, F., Van Roy, V., Eds.; Publications Office of the European Union: Luxembourg, 2018. Available online: https://ec.europa.eu/jrc/sites/default/files/eides_2018.pdf (accessed on 4 June 2021). [CrossRef]
23. Nambisan, S. Digital entrepreneurship: Toward a digital technology perspective of entrepreneurship. *Entrep. Theory Pract.* **2017**, *41*, 1029–1055. [CrossRef]
24. EC. Digital Transformation of European Industry and Enterprises: A Report of the Strategic Policy Forum on Digital Entrepreneurship. 2015. Available online: https://ec.europa.eu/growth/content/report-digital-transformation-european-industry-and-enterprises_en (accessed on 4 June 2021).
25. Kraus, S.; Palmer, C.; Kailer, N.; Kallinger, F.L.; Spitzer, J. Digital entrepreneurship: A research agenda on new business models for the twenty-first century. *Int. J. Entrep. Behav. Res.* **2019**, *25*, 353–375. [CrossRef]
26. Autio, E.; Szerb, L.; Komlósi, É.; Tiszberger, M. *The European Index of Digital Entrepreneurship Systems*; Nepelski, D., Ed.; Publications Office of the European Union: Luxembourg, 2020. Available online: https://joint-research-centre.ec.europa.eu/european-index-digital-entrepreneurship-systems-eides_en (accessed on 4 June 2021). [CrossRef]
27. Bican, P.M.; Brem, A. Digital Business Model, Digital Transformation, Digital Entrepreneurship: Is There A Sustainable "Digital"? *Sustainability* **2020**, *12*, 5239. [CrossRef]
28. Walsh, P.P.; Murphy, E.; Horan, D. The role of science, technology and innovation in the UN 2030 agenda. *Technol. Forecast. Soc. Chang.* **2020**, *154*, 119957. [CrossRef]
29. Jovanović, M.; Dlačić, J.; Okanović, M. Digitalization and society's sustainable development–Measures and implications. *Zb. Rad. Ekon. Fak. U Rijeci: Časopis Za Ekon. Teor. I Praksu* **2018**, *36*, 905–928. [CrossRef]
30. Tilson, D.; Lyytinen, K.; Sørensen, C. Research commentary-digital infrastructures: The missing IS research agenda. *Inf. Syst. Res.* **2010**, *21*, 748–759. [CrossRef]
31. Wen, H.; Lee, C.C.; Song, Z. Digitalization and environment: How does ICT affect enterprise environmental performance? *Environ. Sci. Pollut. Res.* **2021**, *28*, 4826–54841. [CrossRef] [PubMed]
32. Ghobakhloo, M.; Iranmanesh, M. Digital transformation success under Industry 4.0: A strategic guideline for manufacturing SMEs. *J. Manuf. Technol. Manag.* **2021**, *32*, 1533–1556. [CrossRef]
33. Mohamed Hashim, M.A.; Tlemsani, I.; Matthews, R. Higher education strategy in digital transformation. *Educ. Inf. Technol.* **2022**, *27*, 3171–3195. [CrossRef]
34. Satalkina, L.; Steiner, G. Digital Entrepreneurship: A Theory-Based Systematization of Core Performance Indicators. *Sustainability* **2020**, *12*, 4018. [CrossRef]
35. Satalkina, L.; Steiner, G. Digital entrepreneurship and its role in innovation systems: A systematic literature review as a basis for future research avenues for sustainable transitions. *Sustainability* **2020**, *12*, 2764. [CrossRef]
36. Block, J.H.; Fisch, C.O.; van Praag, M. The Schumpeterian entrepreneur: A review of the empirical evidence on the antecedents, behaviour and consequences of innovative entrepreneurship. *Ind. Innov.* **2017**, *24*, 61–95. [CrossRef]
37. Urbano, D.; Aparicio, S.; Audretsch, D. Twenty-five years of research on institutions, entrepreneurship, and economic growth: What has been learned? *Small Bus. Econ.* **2019**, *53*, 21–49. [CrossRef]
38. Szabo, Z.K.; Herman, E. Innovative entrepreneurship for economic development in EU. *Procedia Econ. Financ.* **2012**, *3*, 268–275. [CrossRef]
39. Audretsch, D.B.; Peña-Legazkue, I. Entrepreneurial activity and regional competitiveness: An introduction to the special issue. *Small Bus. Econ.* **2012**, *39*, 531–537. [CrossRef]

40. Henrekson, M.; Sanandaji, T. Small Business Activity Does Not Measure Entrepreneurship. *Proc. Natl. Acad. Sci. USA* **2014**, *111*, 1760–1765. [CrossRef] [PubMed]
41. Szabo, Z.K.; Soltes, M.; Herman, E. Innovative capacity & performance of transition economies: Comparative study at the level of enterprises. *Econ. Manag.* **2013**, *16*, 52–68.
42. Jafari-Sadeghi, V. The motivational factors of business venturing: Opportunity versus necessity? A gendered perspective on European countries. *J. Bus. Res.* **2020**, *113*, 279–289. [CrossRef]
43. Ivanović-Đukić, M.; Stevanović, T.; Rađenović, T. Does digitalization affect the contribution of entrepreneurship to economic growth? *Zb. Rad. Ekon. Fak. U Rijeci: Časopis Za Ekon. Teor. I Praksu* **2019**, *37*, 653–679. [CrossRef]
44. Sussan, F.; Acs, Z.J. The digital entrepreneurial ecosystem. *Small Bus. Econ.* **2017**, *49*, 55–73. [CrossRef]
45. Mrożewski, M.; Kratzer, J. Entrepreneurship and country-level innovation: Investigating the role of entrepreneurial opportunities. *J. Technol. Transf.* **2017**, *42*, 1125–1142. [CrossRef]
46. Dhahri, S.; Slimani, S.; Omri, A. Behavioral entrepreneurship for achieving the sustainable development goals. *Technol. Forecast. Soc. Chang.* **2021**, *165*, 120561. [CrossRef]
47. Baumol, W.J. Entrepreneurship: Productive, unproductive, and destructive. *J. Polit. Econ.* **1990**, *98*, 893–921. [CrossRef]
48. Baumol, W.J.; Strom, R.J. Entrepreneurship and economic growth. *Strateg. Entrep. J.* **2007**, *1*, 233–237. [CrossRef]
49. Guerrero, M.; Liñán, F.; Cáceres-Carrasco, F.R. The influence of ecosystems on the entrepreneurship process: A comparison across developed and developing economies. *Small Bus. Econ.* **2020**, *57*, 1733–1759. [CrossRef]
50. Gawel, A. International Trade in the High-Tech Sector—Support or Obstacle to Start-Up Processes at the Macro Level in European Union Countries? *J. Theor. Appl. Electron. Commer. Res.* **2021**, *16*, 1877–1892. [CrossRef]
51. Cao, Z.; Shi, X. A systematic literature review of entrepreneurial ecosystems in advanced and emerging economies. *Small Bus. Econ.* **2021**, *57*, 75–110. [CrossRef]
52. The Global Entrepreneurship and Development Institute (GEDI). The Global Entrepreneurship Index 2019. Available online: https://thegedi.org/wp-content/uploads/2020/01/GEI_2019_Final-1.pdf (accessed on 20 March 2021).
53. Zhao, F. Exploring the synergy between entrepreneurship and innovation. *Int. J. Entrep. Behav. Res.* **2005**, *11*, 25–41. [CrossRef]
54. Ionescu, G.H.; Firoiu, D.; Pîrvu, R.; Enescu, M.; Rădoi, M.I.; Cojocaru, T.M. The potential for innovation and entrepreneurship in EU countries in the context of sustainable development. *Sustainability* **2020**, *12*, 7250. [CrossRef]
55. Schumpeter, J. *The Theory of Economic Development*; Harvard University Press: Cambridge, MA, USA, 1934.
56. Herman, E. Innovation and entrepreneurship for competitiveness in the EU: An empirical analysis. In Proceedings of the International Conference on Business Excellence, Bucharest, Romania, 22–23 March 2018; pp. 425–435. [CrossRef]
57. Dagilienė, L.; Bruneckienė, J.; Jucevičius, R.; Lukauskas, M. Exploring smart economic development and competitiveness in Central and Eastern European countries. *Compet. Rev.* **2020**, *30*, 485–505. [CrossRef]
58. Si, S.; Zahra, S.A.; Wu, X.; Jeng, D.J.F. Disruptive innovation and entrepreneurship in emerging economics. *J. Eng. Technol. Manag.* **2020**, *58*, 101601. [CrossRef]
59. World Economic Forum (WEF). The Global Competitiveness Report 2019. Available online: https://www3.weforum.org/docs/WEF_TheGlobalCompetitivenessReport2019.pdf (accessed on 4 March 2021).
60. European Union (EU). *Digital Transformation Scoreboard 2018: EU Businesses Go Digital: Opportunities, Outcomes and Uptake*; Publications Office of the European Union: Luxembourg, 2018.
61. Soluk, J.; Kammerlander, N.; Darwin, S. Digital entrepreneurship in developing countries: The role of institutional voids. *Technol. Forecast. Soc. Chang.* **2021**, *170*, 120876. Available online: https://www.sciencedirect.com/science/article/pii/S0040162521003085 (accessed on 4 October 2021). [CrossRef]
62. Von Briel, F.; Davidsson, P.; Recker, J. Digital technologies as external enablers of new venture creation in the IT hardware sector. *Entrep. Theory Pract.* **2018**, *42*, 47–69. [CrossRef]
63. Abubakre, M.; Faik, I.; Mkansi, M. Digital entrepreneurship and indigenous value systems: An Ubuntu perspective. *Inf. Syst. J.* **2021**. Available online: https://onlinelibrary.wiley.com/doi/10.1111/isj.12343 (accessed on 4 August 2021). [CrossRef]
64. The World Bank. *World Development Report 2016: Digital Dividends*; World Bank Publications: Washington DC, USA, 2016. Available online: https://www.worldbank.org/en/publication/wdr2016 (accessed on 4 June 2021).
65. Martin-Rojas, R.; Garcia-Morales, V.J.; Gonzalez-Alvarez, N. Technological antecedents of entrepreneurship and its consequences for organizational performance. *Technol. Forecast. Soc. Chang.* **2019**, *147*, 22–35. [CrossRef]
66. Du, W.; Pan, S.L.; Zhou, N.; Ouyang, T. From a marketplace of electronics to a digital entrepreneurial ecosystem (DEE): The emergence of a meta-organization in Zhongguancun, China. *Inf. Syst. J.* **2018**, *28*, 1158–1175. [CrossRef]
67. Elia, G.; Margherita, A.; Passiante, G. Digital entrepreneurship ecosystem: How digital technologies and collective intelligence are reshaping the entrepreneurial process. *Technol. Forecast. Soc. Chang.* **2020**, *150*, 119791. [CrossRef]
68. Nambisan, S.; Siegel, D.; Kenney, M. On open innovation, platforms, and entrepreneurship. *Strateg. Entrep. J.* **2018**, *12*, 354–368. [CrossRef]
69. Samara, G.; Terzian, J. Challenges and Opportunities for Digital Entrepreneurship in Developing Countries. In *Digital Entrepreneurship*; Soltanifar, M., Hughes, M., Göcke, L., Eds.; Springer: Cham, Switzerland, 2021; pp. 283–302. [CrossRef]
70. Shamsrizi, M.; Pakura, A.; Wiechers, J.; Pakura, S.; Dauster, D.V. Digital Entrepreneurship for the "Decade of Action". In *Digital Entrepreneurship*; Soltanifar, M., Hughes, M., Göcke, L., Eds.; Springer: Cham, Switzerland, 2021; pp. 303–327. [CrossRef]

71. Hull, C.E.K.; Hung, Y.T.C.; Hair, N.; Perotti, V. Taking advantage of digital opportunities: A typology of digital entrepreneurship. *Int. J. Netw. Virtual Organ.* **2007**, *4*, 290–303. [CrossRef]
72. Le Dinh, T.; Vu, M.C.; Ayayi, A. Towards a living lab for promoting the digital entrepreneurship process. *Int. J. Entrep.* **2018**, *22*, 1–17. Available online: https://www.abacademies.org/articles/Towards-a-living-lab-for-promoting-the-digital-entreprenp-process-%2022-1-130.pdf (accessed on 20 September 2021).
73. Song, A.K. The Digital Entrepreneurial Ecosystem—A critique and reconfiguration. *Small Bus. Econ.* **2019**, *53*, 569–590. [CrossRef]
74. Sahut, J.M.; Iandoli, L.; Teulon, F. The age of digital entrepreneurship. *Small Bus. Econ.* **2021**, *56*, 1159–1169. [CrossRef]
75. Munoz, P.; Cohen, B. Sustainable entrepreneurship research: Taking stock and looking ahead. *Bus. Strateg. Environ.* **2018**, *27*, 300–322. [CrossRef]
76. York, J.G.; Venkataraman, S. The entrepreneur–environment nexus: Uncertainty, innovation, and allocation. *J. Bus. Ventur.* **2010**, *25*, 449–463. [CrossRef]
77. Si, S.; Ahlstrom, D.; Wei, J.; Cullen, J. Business, entrepreneurship and innovation toward poverty reduction. *Entrep. Reg. Dev.* **2020**, *32*, 1–20. [CrossRef]
78. Si, S.; Yu, X.; Wu, A.; Chen, S.; Su, Y. Entrepreneurship and poverty reduction: A case study of China. *Asia Pac. J. Manag.* **2015**, *32*, 119–143. [CrossRef]
79. Youssef, A.B.; Boubaker, S.; Omri, A. Entrepreneurship and sustainability: The need for innovative and institutional solutions. *Technol. Forecast. Soc. Chang.* **2018**, *129*, 232–241. [CrossRef]
80. European Union. *Sustainable Development in the European Union Monitoring Report on Progress towards the SDGs in an EU Context*; Publications Office of the European Union: Luxembourg, 2021. Available online: https://ec.europa.eu/eurostat/documents/3217494/12878705/KS-03-21-096-EN-N.pdf/8f9812e6-1aaa-7823-928f-03d8dd74df4f?t=1623741433852 (accessed on 20 September 2021). [CrossRef]
81. Omri, A. Technological innovation and sustainable development: Does the stage of development matter? *Environ. Impact Assess. Rev.* **2020**, *83*, 106398. [CrossRef]
82. Portulans Institute. *The Network Readiness Index 2020: Accelerating Digital Transformation in a Post-COVID Global Economy*; Dutta, S., Lanvin, B., Eds.; Portulans institute: Washington, WA, USA, 2020. Available online: https://networkreadinessindex.org/wp-content/uploads/2020/10/NRI-2020-Final-Report-October2020.pdf (accessed on 4 March 2021).
83. George, G.; Merrill, R.K.; Schillebeeckx, S.J. Digital sustainability and entrepreneurship: How digital innovations are helping tackle climate change and sustainable development. *Entrep. Theory Pract.* **2021**, *45*, 999–1027. [CrossRef]
84. EC. A Digital Single Market Strategy for Europe. COM (2015) 192 Final. 2015. Available online: https://eur-lex.europa.eu/legalcontent/RO/TXT/PDF/?uri=CELEX:52015DC0192&from=EN (accessed on 4 March 2021).
85. EC. Digital Economy and Society Index (DESI) 2020: Thematic Chapters. 2020. Available online: https://eufordigital.eu/wp-content/uploads/2020/06/DESI2020Thematicchapters-FullEuropeanAnalysis.pdf (accessed on 4 March 2021).
86. Eurostat Database. 2021. Available online: http://ec.europa.eu/eurostat/data/database (accessed on 20 March 2021).
87. EC. European Innovation Scoreboard (EIS) 2021—Database. Available online: https://ec.europa.eu/docsroom/documents/46934 (accessed on 4 March 2021).
88. The Global Entrepreneurship and Development Institute (GEDI) Datasets. Global Entrepreneurship Research Data. 2021. Available online: https://thegedi.org/datasets/ (accessed on 4 March 2021).
89. EC. Datasets. In *Digital Economy and Society Index (DESI)*; 2021. Available online: https://digital-agenda-data.eu/datasets/desi/visualizations (accessed on 4 March 2021).
90. WEF. The Global Competitiveness Report. 2018. Available online: https://www3.weforum.org/docs/GCR2018/05FullReport/TheGlobalCompetitivenessReport2018.pdf (accessed on 4 March 2021).
91. Portulans Institute. *The Network Readiness Index 2019: Towards a Future-Ready Society*; Dutta, S., Lanvin, B., Eds.; Portulans institute: Washington, WA, USA, 2020. Available online: https://networkreadinessindex.org/2019/wp-content/uploads/2020/03/The-Network-Readiness-Index-2019-New-version-March-2020.pdf (accessed on 4 March 2021).
92. Sustainable Development Solutions Network (SDSN) and Institute for European Environmental Policy (IEEP). *The 2020 Europe Sustainable Development Report: Meeting the Sustainable Development Goals in the Face of the COVID-19 Pandemic*; Sustainable Development Solutions Network and Institute for European Environmental Policy: Paris, France; Brussels, Belgium, 2020. Available online: https://www.sdgindex.org/reports/europe-sustainable-development-report-2020/ (accessed on 4 June 2021).
93. SDSN and IEEP. *The 2019 Europe Sustainable Development Report. Towards a Strategy for Achieving the Sustainable Development Goals in the European Union*; Sustainable Development Solutions Network and Institute for European Environmental Policy: Paris, France; Brussels, Belgium, 2019. Available online: https://www.sdgindex.org/reports/2019-europe-sustainable-development-report/ (accessed on 4 March 2021).
94. GEDI. The Global Entrepreneurship Index 2018. Available online: https://thegedi.org/wp-content/uploads/dlm_uploads/2017/11/GEI-2018-1.pdf (accessed on 20 March 2021).
95. EC. European Innovation Scoreboard 2021—Methodology Report. Available online: https://ec.europa.eu/docsroom/documents/45971 (accessed on 20 March 2021).
96. Lafortune, G.; Schmidt-Traub, G. Study on Exposing EU Policy Gaps to Address the Sustainable Development Goals European Union, SDSN 2018. Available online: https://www.eesc.europa.eu/sites/default/files/files/qe-02-19-009-en-n.pdf (accessed on 20 March 2021).

97. Agresti, A. *An Introduction to Categorical Data Analysis*, 2nd ed.; John Wiley & Sons, Inc.: Hoboken, NJ, USA, 2007.
98. Kovács, T.Z.; Bittner, B.; Huzsvai, L.; Nábrádi, A. Convergence and the Matthew Effect in the European Union Based on the DESI Index. *Mathematics* **2022**, *10*, 613. [CrossRef]
99. Landau, S.; Everitt, B.S. *A Handbook of Statistical Analyses Using SPSS*; Chapman & Hall/CRC Press LLC: Boca Raton, FL, USA, 2004.
100. Jolliffe, I. *Principal Component Analysis*, 2nd ed.; Springer: New York, NY, USA, 2002.
101. Herman, E. Working poverty in the European Union and its main determinants: An empirical analysis. *Inz. Ekon.* **2014**, *25*, 427–436. [CrossRef]
102. Everit, B.S.; Landau, S.; Leese, M.; Stahl, D. *Cluster Analysis*, 5th ed.; Willey: London, UK, 2011.
103. Georgescu, M.A.; Herman, E. Productive employment for inclusive and sustainable development in European Union countries: A multivariate analysis. *Sustainability* **2019**, *11*, 1771. [CrossRef]
104. The World Bank. Doing Business 2020: Comparing Business Regulation in 190 Economies. 2020. Available online: https://documents1.worldbank.org/curated/en/688761571934946384/pdf/Doing-Business-2020-Comparing-Business-Regulation-in-190-Economies.pdf (accessed on 20 March 2021).
105. Raut, J.; Ćelić, Đ.; Dudić, B.; Ćulibrk, J.; Stefanović, D. Instruments and Methods for Identifying Indicators of a Digital Entrepreneurial System. *Mathematics* **2021**, *9*, 2151. [CrossRef]
106. EC. Shaping the Digital Transformation in Europe. In *McKinsey & Company Report*; 2020. Available online: https://digital-strategy.ec.europa.eu/en/library/shaping-digital-transformation-europe (accessed on 20 March 2022).
107. The United Nations Development Programme 2018. In *Development 4.0: Opportunities and Challenges for Accelerating Progress towards the Sustainable Development Goals in Asia and the Pacific*; 2018. Available online: https://www.asia-pacific.undp.org/content/rbap/en/home/library/sustainable-development/Asia-Pacific-Development-40.html (accessed on 20 March 2022).

mathematics

Article

Computer Vision Algorithms, Remote Sensing Data Fusion Techniques, and Mapping and Navigation Tools in the Industry 4.0-Based Slovak Automotive Sector

Marek Nagy [1,*] and George Lăzăroiu [2,*]

[1] Faculty of Operation and Economics of Transport and Communications, University of Zilina, Univerzitna 1, 01026 Zilina, Slovakia
[2] Department of Economic Sciences, Spiru Haret University, 030045 Bucharest, Romania
* Correspondence: nagy16@stud.uniza.sk (M.N.); george.lazaroiu@spiruharet.ro (G.L.)

Abstract: The objectives of this paper, and the novelty brought to the topic of the Industry 4.0 manufacturing systems, are related to the integration of computer vision algorithms, remote sensing data fusion techniques, and mapping and navigation tools in the Slovak automotive sector. We conducted a thorough examination of Industry 4.0-based value and supply chains, clarifying how cyber-physical production systems operate in relation to collision avoidance technologies, environment mapping algorithms, and mobility simulation tools in network connectivity systems through vehicle navigation data. The Citroen C3 and Peugeot 208 automobiles are two examples of high-tech products whose worldwide value and supply chain development trends were examined in this study by determining countries and their contributions to production. The fundamental components of the research—statistical analysis and visual analysis—were utilized in conjunction with a variety of syntheses, comparisons, and analytical methodologies. A case study was developed using PSA Group SVK data. The graphical analysis revealed that Slovakia offers the second-highest added value to the chosen items, but it also highlighted the country's slow-growing research and development (R&D) infrastructure, which could lead to a subsequent loss of investment and business as usual. Slovakia can generate better export added value by optimizing Industry 4.0-based manufacturing systems in the automotive sector.

Keywords: Industry 4.0; value-added; global value chain; export; globalization; innovation; digitization; automotive industry

MSC: 91-11

1. Introduction

The climate of the large-scale economy is changing fundamentally as a result of globalization, leaving behind a world when hurdles to foreign trade and restricted investment transfers caused by geography, time zones, language, or national variations in governmental regulation, distinct cultures, or dissimilar business systems kept national economies largely isolated from one another [1]. As we move into this new era, trade restrictions are being reduced and, in some cases, even eliminated entirely. Due to advancements in transportation and telecommunication technologies, the concept of distance is becoming less significant [2] and global material culture tends to unify. An interdependent, integrated global system brings together national economies [3].

All facets of society are affected directly or indirectly by globalization, but an examination of its underlying causes also reveals two features of its growth that stand out as particularly significant:

1. Innovations in technology, the growth of information systems, and advancements in manufacturing processes make up the first dimension. The physical foundations and tools of globalization are developed in this way.
2. A broad range of social and economic structures enable the configuration of globalization processes [4].

The culmination of all prior development processes, including the deepening of internationalization and integration and the expansion of interdependence and transnationality, as well as large-scale specialization and cooperation, has led to globalization as a determining factor in the growth of national economies [5]. The most recent advancements in science and technology have intensified this tendency to the point that a new structure for the economy is required in order to deploy them effectively.

A wide range of concerns relating to globalization are discussed. International trade is not complete without the presence of large-scale industry and value chains, which give developing nations the chance to become more integrated into it, fight poverty, and open up new employment, production, and innovation opportunities [6]. Globalization has favorable consequences on ecological cooperation, working conditions, and sustainable economic growth. On the other hand, though, globalization's extremely complex nature, lack of transparency, and weakening of commitments can result in an increase in violence, an escalated risk of political violations, disturbances of legal and environmental regulations, and an intensification of tax fraud [7].

The COVID-19 pandemic, the protectionism of some nations, the rising stability of developing countries, the expansion and concurrent decrease of commercial services, and the general dynamics of the international economy are all trends that are currently having an impact on the global industry [8]. However, the beneficial effects exceed the negative ones, thus international trade, global value chains, and the adoption of cutting-edge economic trends are necessary [9]. The emergence of megatrends in terms of their weight and impact on global commerce leads to the growing importance of international trade and its long-term developments. We distinguish six related factors [10]: (i) interdependence (growing mutual dependence); (ii) integration (regional and global); (iii) transnationality (transcending national frameworks); (iv) scientific and technical progress; (v) adaptation and cooperation of basic entities of international trade; (vi) threats of global problems.

Specifically, because of the effects of globalization, extensive technical–technological innovations and the application of scientific–technical knowledge in all spheres of business activity have emerged as significant and vital drivers of the dynamic growth of international trade [11]. The positions of the major transnational corporations (TNCs) and entire nations are shifting quickly due to the explosive growth in the proportion of goods and services with a high level of added value (high-tech) in global trade [12]. Such simultaneous processes—advancing the existing technologies, the introduction of radical innovations that qualitatively alter the mode of production, the beginning of big data-driven discoveries, and the realization of scientific and technical knowledge—evidence the aforementioned complexity of scientific and technological progress [13]. The emergence of Industry 4.0 technologies lowers entry barriers and virtualizes the global value chain [14]. More and more edge service providers can now enter the market thanks to this innovation surge.

The main objective of this paper is to analyze the development of Industry 4.0-based manufacturing systems, a current trend in digitalization and associated production automation, as a means of enhancing sustainable business performance in particular Slovak national circumstances across the automotive industry sector in terms of computer vision algorithms, remote sensing data fusion techniques, and mapping and navigation tools. The findings of this study are significant on both a sectoral and a national level because Slovakia's automobile industry shapes the national economy considerably. The study's main contribution emphasizes the importance of the automotive sector and the rise in added value, in addition to the influence of Industry 4.0 on the international trade structure and the changes across the global value chains. The analysis covers mainly PSA Group SVK (Slovakia), integrating data across the EU, and emphasizing how Industry 4.0-based manufacturing systems deploy visual perception and remote sensing technologies, route detection and object localization algorithms, real-world connected vehicle navigation data, big geospatial data analytics, and trajectory planning and mobility simulation tools throughout the automotive sector and networked transport systems.

The paper is divided as follows: The Literature Review introduces our research into the body of knowledge, which ranges from global megatrends to the distinct Industry 4.0 revolution. The Material and Methods section explores the data and techniques that are currently available to help determine whether these phenomena apply to the Slovak industry. The subsequent section analyses the findings before outlining a discussion of the many implications of Industry 4.0 in relation to raising Slovakia's added value. The study's conclusions are that data visualization tools, cognitive data fusion techniques, and deep learning-based computer vision algorithms are pivotal in digital twin-based cyber-physical production systems in the Slovak automotive sector. Prospective future visions are suggested as regards how cloud computing and remote sensing technologies, immersive visualization tools, and cyber-physical production systems can be operational in virtual enterprises and immersive 3D environments by use of virtual modeling and simulation tools, digital twin modeling, and neural network algorithms in the Slovak automotive sector.

2. Literature Review

Industry 4.0, which is founded on the Internet of Things sensing networks, cyber-physical system-based manufacturing, cognitive automation, and deep learning-assisted smart process planning in the Slovak automotive sector, is the primary focus of the research. However, it is crucial to recall and describe the large-scale economy of the globalization age, which was marked by transnationality, Foreign Direct Investment (FDI), global value chains, and exports, in order to grasp this issue theoretically. These subtopics are covered in the first stage of the research since they are closely related to and contribute to the growth of the main debated issue.

Gereffi [15] originally proposed the idea of the global value chain (GVC) using the specific example of the apparel sector, integrating the challenges and competitiveness of businesses. GVC represents a series of social activities that give rise to the value of its goods and services [16,17]. According to Mugge [18], the added value can be assessed quite thoroughly in the economies that generate it through participation. There is a significant shift in the removal of sporadic restrictions to international trade as being related to globalization and its pressures. An ideal input–output ratio is the aim and final result of global value chains according to Ye et al. [19]. The importance of international trade and its continuous dynamics have been influenced by the advent of several factors, which are usually referred to as megatrends in terms of their weight and impact on large-scale commerce [20]. One of these factors is transnationality, which has given rise to new kinds of agreements on intra-company commerce and inter-corporate collaboration [21]. Transnational corporations, associated with FDI, are an accurate representation of every transaction made between the direct investor and the business. Exports of FDI either replace transnationality or strengthen it [22]. FDI demonstrates a long-term interest by resident entities in one economy by controlling a resident firm in another country [23]. Scientific advancement is the most significant accelerator for the dynamic rise of international trade [24]. Because of the growth of artificial intelligence-based decision-making algorithms, the Internet of Things sensing networks, and cyber-physical production systems, which decrease entry barriers for a growing volume of marginal service providers through outsourcing or offshore sourcing [25], the value chain is being virtualized, integrating machine and deep learning technologies, digital twin algorithms, and spatial data visualization tools in the Slovak automotive sector. According to Minarik et al. [4], the history of globalization shows that two characteristics of its development are decisive: substantial changes in the technological sphere of production and the variety of social and economic forms. Based on [26,27], the first portion of the paper examines the four industrial revolutions and their significant contributions to globalization.

Kagermann and Wahlster (2011) introduced the idea of Industry 4.0 designed as a strategic initiative for the creation of groundbreaking manufacturing systems, with the goal of boosting productivity and efficiency across the national industry [28]. According to Zabojnik [29], Industry 4.0 was introduced during the third revolution, utilizing the Internet in all facets of industrial production and enabling real-time machine–machine, man–machine, and man–man communication [30]. Because the Internet is presently the largest network in existence, self-configurable networks with the ability for autonomous configuration are mentioned by Clayton and Kral [31] as a phenomenon of the twenty-first century. In addition, Clayton and Kral [31] forecast an Industry 4.0-based massive growth of the Internet across all spheres of human activity as well as the linking of the physical and virtual worlds over a period of 10 to 30 years. Industry 4.0, according to Yang and Gu [32], is the integration of Internet technologies into industrial production networks that will result in considerable output improvements, real-time shop floor connectivity, and a linkage between production and sales of manufactured items. Hermann et al. [33] examine the possibility of achieving the necessary state of networking based on real-time data from both the physical world and the virtual, i.e., gathered online content.

Industry 4.0, a term used by Klingenberg et al. [34] to describe horizontally connected activities within the value stream, is closely related to digitization. Industry 4.0 and the Internet of Things are included in this flow throughout the manufacturing and usage phases, with digitization serving as a connector. Industry 4.0 integration, according to Rogers and Zvarikova [35], would boost productivity by 6% to 8% annually in global trade. Ruttimann and Stickli [36] underline the growing need to put data-driven technology networking into practice and achieve full connectivity. The goal of Slovakia's smart industry is to integrate cyber-physical production technologies into all facets of the country's economy [37]. Kovacova and Lewis [38] highlight the difficult-to-define topic of Industry 4.0, which they describe as the transition of production from discrete automated units to a completely automated and continually optimized production environment. This is accomplished by building groundbreaking worldwide networks based on the linking of manufacturing facilities into cyber-physical production systems, or smart factories, that function as production facilities. According to Mehmann and Teuteberg [39], the digital value chain represents a future in which product quality will increase and delivery times will decrease. Industry 4.0 is supported by three pillars: traditional industry, digital technology, and the Internet, say Galbraith and Podhorska [40]. Industry 4.0 will primarily manifest itself in the fields of communication networks and artificial intelligence data-driven Internet of Things systems. Sony [41] reports on both the possible advantages and threats that Industry 4.0 may bring, noting that the former outweighs the latter and that the onset of the Internet of Things-based real-time production logistics is inevitable.

In addition to improvements in the economy as a whole, Dalenogare et al. [42] also note optimizations in the social sectors of society. The linked supply chain will be impacted by digitalization, which will lower costs and improve end-to-end process management through the application of industrial artificial intelligence. When approaching Industry 4.0, one should consider how data might add value and what function each technological advancement serves, that is, its originality and worth [43]. There is a clear connection between Industry 4.0 technology and real-time big data analytics, a major driver of this transformation. Research by Svabova et al. [44] addressing the development of the role of employees throughout automated production processes describes the cutting-edge items produced by harnessing intelligent manufacturing. The value chain will be considerably improved by digitization and robotic wireless sensor networks, which will boost productivity, save costs, and foster greater innovation and cooperation through real-time big data analytics [45]. Four fundamental characteristics—vertical connectivity of intelligent production systems, Internet of Things-based real-time production logistics, digitized mass production, and interconnected virtual services in cyber-physical system-based smart factories—are used by Said et al. (2021) to describe and explain Industry 4.0-based manufacturing systems. Networking of clients and business partners will then be configured by the horizontal integration of global networks in the Slovak automotive sector.

Industry 4.0 is based on the direct communication and collaboration between people, machines, equipment, logistics systems, and goods, claim Lawrence and Durana [46]. Zavadska and Zavadsky [47] emphasize the value of digital data and the promptness of management in developing corporate strategies through business process optimization. Dynamic transformation brought about by Industry 4.0-based manufacturing systems transcends real-time sensor networks. Globalization may propel national economies to groundbreaking innovations and hence greatly raise the degree of creativity in a nation [48].

Lazaroiu and Harrison [49] assert that the Internet of Things (IoT) and smart sensors [50] are key components of Industry 4.0 manufacturing systems, by integrating cognitive decision-making algorithms, cyber-physical production networks, and sustainable organizational performance. Using mathematical optimization models based on multi-correlation dependencies, Zhong et al. [51] provide a wide range of options for real-time data processing. Belhadi et al. [52] propose that collaboration between supply chain stakeholders will be important to overcome the difficulties of the COVID-19 crisis and expedite the deployment of digital technology. In order for businesses to have greater profits while maximizing their productivity and competitiveness by leveraging product decision-making information systems, digitization necessitates investing in sufficient steps to adapt to digital transformation. From consumer perspectives, more and better services will be available, increasing satisfaction with suitable services in the Slovak automotive sector.

The upshot of the expanding division of labor constitutes the creation and growth of GVC. Production activities are highly fragmented and distributed among various economies in the world [53]. Each nation concentrates on the distinct stages of the industrial process, integrating competitive advantages. The highest added value is cumulated by countries that engage in providing cutting-edge services or research and development (R&D) activities [7]. Despite the fact that value chains seem to be global, some production phases are articulated in different parts of the world. More complex and technologically demanding operations are typically implemented in more developed nations, while intermediate consumption and finalization (assembly) are performed in less advanced economies (Figure 1).

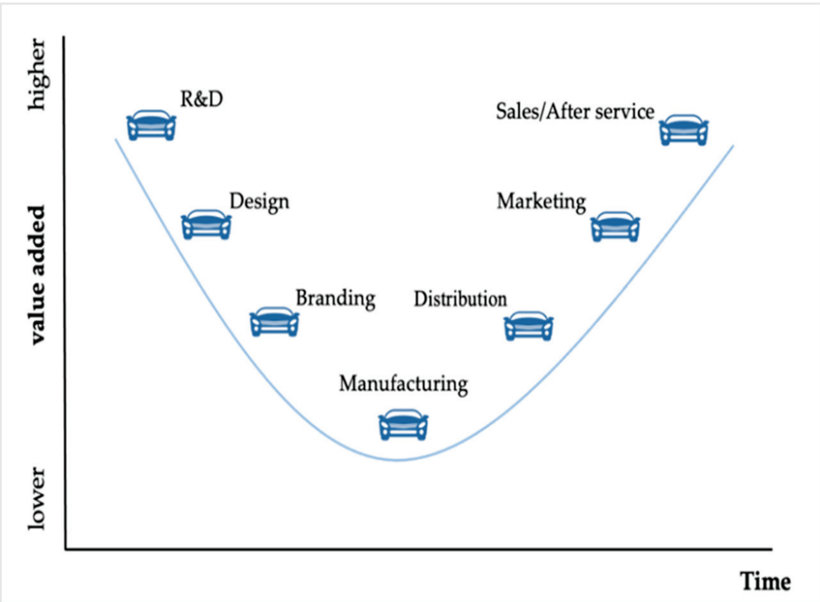

Figure 1. Value-added in global value chain (individual processes). Source: Authors' compilation.

The curve illustrates potential avenues for increasing production with increased added value. Such regions primarily appear at the start and finish of the value chain [54]. How to improve the added value of the automobile sector can be determined using straightforward findings regarding the optimized competitive advantage [9,10]. In Japan, having a skilled worker pool represents a source of creativity. Thus, improving the tactics in global value chains may lead to producing more added value [55]. Revolutionary developments and enhanced process efficiency are hallmarks of operational advancements (such as in Guatemalan handicrafts, which compete with goods from Asia). Product upgrading refers to changes made to the product line to increase value (the coffee industry, for example). Utilizing processes that provide more added value, such as marketing, sales, services, design, and R&D, results in functional upgrading (for instance, Mexico and its techniques for producing and exporting jeans). Interchain upgrading refers to adjustments to a company's manufacturing processes that provide opportunities to new global markets (e.g., Taiwan and its approach to computer production). Setting up and shifting global value chains can bring about uncertain outcomes. Outstanding leaders in the industry protect their patent rights and, as a result, maintain their positions while steadily growing their market share [56].

Industry 4.0 is now underway and attempts to enhance innovation, technology, ecological policy, and education [57]. This is a result of the Internet's explosive expansion and related technological improvements. These are the fundamental elements of Industry 4.0: (i) vertical connection (integrating ground-breaking manufacturing systems, logistics, production, and marketing); (ii) horizontal integration (networking clients and business partners, advanced business models, and large-scale production networks); (iii) technology application (throughout the entire product life cycle); (iv) the market's adoption of exponential technologies is accelerated by the decline in their operational costs. Digital twin-based product development, autonomous manufacturing systems, virtual reality modeling tools, and spatial data visualization techniques are among the areas where Industry 4.0 can be developed in the Slovak automotive sector. Depending on their primary objective, Industry 4.0 technologies can be separated into two distinct tiers. The front-end technologies of Industry 4.0, which take into account the transformation of manufacturing activities based on new technologies (smart manufacturing) and the way that goods are offered (smart products), will be placed in the core of the framework [58]. Additionally included are the mode of delivery (smart supply chain) and processing (smart working) [59] (Figure 2).

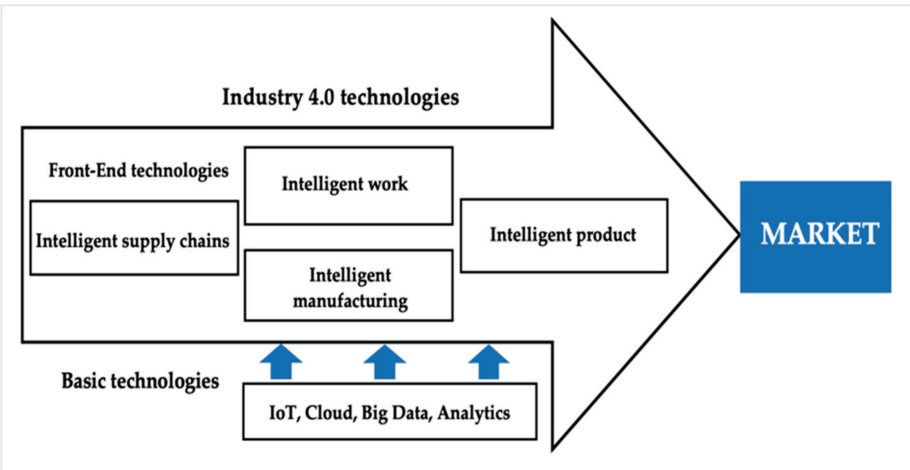

Figure 2. Scheme of Industry 4.0 technologies. Source: Authors' compilation.

Standardization is one of the difficulties in putting Industry 4.0 into practice. It appears almost obligatory in this regard to subject all standards, including internal ones, to the needs and international criteria created in collaboration with significant global actors across international platforms [60]. Above all, it is important to determine if implementing

Industry 4.0 would provide the EU with a competitive advantage in international markets or whether doing so will help it maintain its current position by incorporating sustainable manufacturing and the Internet of Manufacturing Things. Or, at worst, if rapidly expanding economies such as China will inevitably take over as the industrial leader as a result of the global dispersion of technology through multinational corporations [61]. Analyzing how this revolution can result in Slovakia (or in a similar technologically advanced economy) achieving the desired reputation of a creative nation rather than the status of an assembly workshop is conceivable using the example of value and supply chains in the automotive industry in Slovakia.

3. Materials and Methods

By inspecting the recent literature and data on the Industry 4.0-based automotive sector, we identified the most relevant technologies in terms of operational efficiency and practical outcomes, that is, computer vision algorithms, remote sensing data fusion techniques, and mapping and navigation tools, and checked how they performed in the Slovak car and part manufacturing systems.

The advent of the German automaker Volkswagen AG was associated with the growth of the automotive sector in Slovakia. In the vicinity of Bratislava, the first factory was built in 1998. The second round of investments began in 2003 and ended in 2006 with the establishment of plants by PSA Peugeot Citroen in Trnava and KIA Motors in Zilina. A Jaguar Land Rover facility opened up close to Nitra in 2015. The automotive sector took over as the key driver of Slovakia's economy after the collapse of the markets for the cessation of arms manufacture. Its supply chain was also established gradually, bringing in fresh investments for the economy by integrating computer vision algorithms, remote sensing data fusion techniques, and mapping and navigation tools in the Slovak automotive sector.

To be able to meet the main aim of the study, secondary data analysis was employed to accomplish the paper's main research goal. The primary sources of information in our study were official documents, including books, websites, journal articles, internal records, and government publications [62].

Currently, the major portion of Gross Domestic Product (GDP) comes from the automobile industry (roughly 12%). The most recent advancements in artificial intelligence-based decision-making algorithms, traffic flow prediction tools, and collision avoidance technologies are associated with the arrival of the Swedish automaker Volvo, which will establish itself in the east of Slovakia in 2025 and produce only electric vehicles. The Slovak automobile industry's commodity export structure is mostly composed of passenger cars (HS 8703), their components (HS 8708), and bodywork (HS 8707) [18].

From the perspective of the EU, Slovakia is somewhat dependent on the exports of Germany. Its export performance is over 40%, representing 1688 different categories of exported goods and services. Additionally, the German market is concentrated on more popular items with large added values, particularly automobiles (e.g., VW, BMW, and Mercedes Benz). These operational sectors have a considerable impact on Industry 4.0-based manufacturing systems, where Germany is developing its vision particularly in the engineering and automotive sectors. The establishment of a new automaker, Tesla, is introducing a completely new way of manufacturing that will be able to produce up to 700,000 units annually. This is significant information for the German market as it entails casting every component and developing a more considerate manufacturing process for car batteries. This indicates that the German market and economy is confronting fierce competition (Tesla is already the best-selling electric car in the EU). Slovakian exports have fallen behind in comparison with Germany, Japan, and China in terms of the amount of added value.

Over the course of the 10 years that were monitored (2010–2020, the most recent data available), the value created in the Slovak territory actually decreased (the data mapped in 2019). In the automobile sector, for instance, value-added for goods decreased to EUR 2.8 billion in 2012 (Slovakia's automotive output saw its biggest year-over-year growth in 2007). The highest volume of automobiles ever built in Slovakia was in 2013. When it comes to the production of passenger automobiles per capita, Slovakia now leads globally. Significant electromobility trends have emerged in the previous five years, along with other changes that are frequently associated with Sector 4.0 in the automotive industry. The general tendency is an increase in services, additional value, comfort, and safety brought on by technology and connectivity by integrating deep learning object detection technology, trajectory planning algorithms, and geospatial data visualization tools in the Slovak automotive sector (Figure 3).

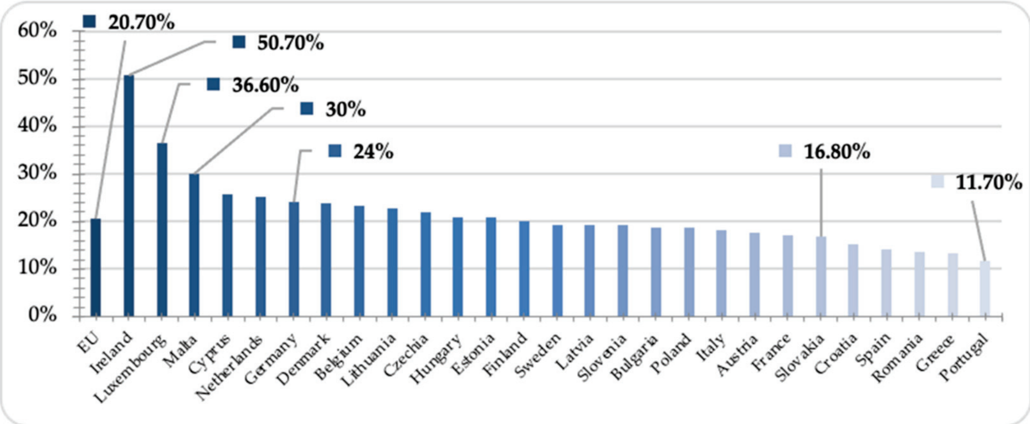

Figure 3. Value-added contribution from exports (EU countries). Source: Authors' compilation according to www.ec.europa.eu (accessed on 4 June 2022).

When examining the exported value-added within the EU, it is crucial to remember and highlight Ireland's tremendous strength (approximately 51%). The majority of Ireland's exports are heavy machinery, while also including chemicals and computer parts. Germany's robust engineering and automotive industries helped it to create 24% of value-added exports. The position of Slovakia among EU countries is 22 (16.8%). As has previously been shown, Slovakia is an assembly-based country based more on manufacturing and less on added value [63]. France, which has not handled the COVID-19 pandemic as well as, say, Germany or China, is in advance of Slovakia. Portugal, a tourist destination where added value predominates mostly in commercial services (tourism), which have decreased by roughly 21% globally as a result of the COVID-19 pandemic, is in last position. As already established, the Slovak economy is primarily propelled by the automotive industry. The electrical and mechanical engineering industries come in second place, receiving up to 30% of all investment projects overall, followed by the automobile industry, including component makers. The car sector, which makes up more than half of the whole industry, deploys driving perception algorithms, big geospatial data analytics, and vehicle routing and navigation systems. Germany and South Korea have made the most investments in Slovakia over the past 18 years, with each contributing approximately 29%, primarily in the automotive and engineering sectors (Germany—VW, South Korea—KIA) [64].

Many analysts and automakers see the current megatrends in the industry as a shift to the ACES model (A—autonomous driving, C—connection, E—electromobility, and S—shared mobility services), by integrating remote sensing technologies, obstacle avoidance algorithms, and mapping and navigation tools [11]. Up to eight of the ten top original car manufacturers anticipate developing autonomous vehicles, shaping the industry's future development through smart traffic planning and analytics, network connectivity systems, and environment mapping algorithms. This is supported by the EU's decision to stop producing cars with internal combustion engines after 2035. According to [65], the cost of purchasing electric cars, the rate of innovation in this area, the capacity development and price of batteries, the infrastructure for electromobility, environmental concerns, and legal and regulatory frameworks will all play a role in the development of electromobility. The level of innovative potential of businesses, education, and worker skills all influence technological and dynamic development. Zabojnik [29] predicts that there will be greater pressures on the workforce in terms of automation and robotization associated with connected vehicle technologies, intelligent transportation planning, and cognitive wireless sensor networks that may reduce employment by up to 30%. Creativity will become a crucial skill with rising demand in Industry 4.0-based manufacturing systems. Artificial intelligence and information technology, which encompass a far greater range of technologies and goods, will call for creativity and the capacity of an individual to deal with a wide range of unpredictable scenarios [66].

Figure 4 shows that Slovakia's gross value-added per employee is in decline (down 40%), whereas the Czech Republic's innovation policy is driving growth (up 48%) and producing high added value [65]. Compared to OECD and EU28 countries, Slovakia requires around twice as much foreign added value (which it imports) for exports. Slovakia's gross exports are made up of imported foreign added value to the tune of almost half (44.8%), dominated by industrial production.

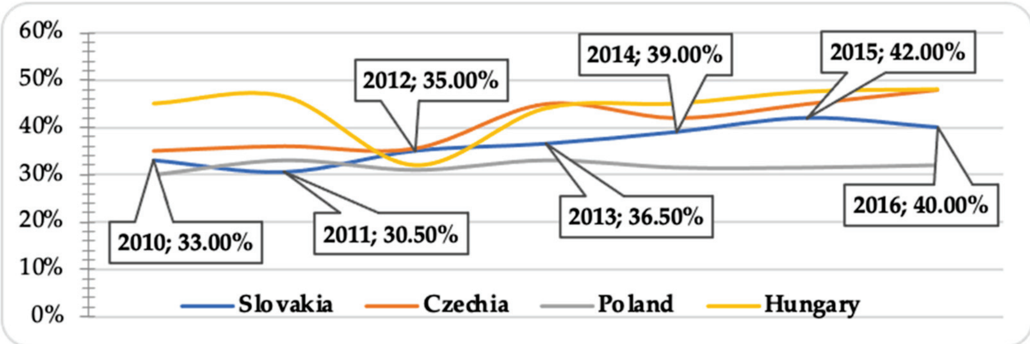

Figure 4. Gross value-added under V4 2010–2017 per employee. Source: Authors' compilation.

Several techniques that are based on actual data received from PSA Group SVK were employed for a more thorough investigation of the added value across the Slovak economy and other countries involved in manufacturing selected automobiles. Statistical and graphical analysis make up the majority of the paper [67]. The work also employs broad analysis and the synthesis of the results that follows. The conclusions of this case study will be generalizable to virtually the entire Slovak sector as we are looking into the added value at PSA Group SVK.

Following the secondary data analysis, which showed how Industry 4.0 was developing under Slovak settings, the following methodological steps were taken to accomplish the paper's main objective:

3. Secondary source research to chart the expansion of Slovakia's automotive industry and the performance of exports from national economies (Sario, economy.gov.sk, datacube, Statistical Office of Slovakia).
4. The main conclusions of the survey, which was collaboratively developed by employees working in Industry 4.0 and digitizing the company's production, were acknowledged. The primary goal was to discover how familiar the chosen company employees were with the concept of Industry 4.0. The second goal was to determine how well-prepared the selected organizations were for the transition to a digital society as a way to increase added value via delivering technical innovations in the form of new products and processes [67]. Building unique methodologies based on the transformation (upgrade) of GVC at the level of process improvements and/or products for the increase of added value was the aim of this research, which was conducted from 1 January to 30 April 2022.
5. Researching and analyzing current automotive industry developments and the impact of Industry 4.0 on the Slovak automotive industry.

4. Results and Discussion

The study's goals included assessing Industry 4.0′s present status in the automotive sector, finding opportunities for value-added development, and spotting existing and emerging trends in driving perception algorithms, vehicle navigation technologies, and remote sensing data fusion techniques. The following phase was conducting a case study to determine whether the automobile sector was ready to implement the Industry 4.0 strategy, which would significantly increase the added value of its product line in future exports by applying industrial big data analytics, robotic wireless sensor networks, and product decision-making information systems in the Slovak automotive sector.

Two case study parts make up this section. The first section displays the PSA Group SVK's supplier structure, including the most significant PSA suppliers and their role in generating more added value. The second section depicts how Slovakia's added value changed within this business before and after Industry 4.0 trends were implemented.

4.1. PSA Company's Supplier Structure (Analysis of Added Value)

In this section of the case study, the aim is to analyze and rank PSA Group SVK's supply and value chains for the Citroen C3 and Peugeot 208 cars for the available monitored years, as well as to visually depict the amount of added value Slovakia created during the production of these cars (Figure 5).

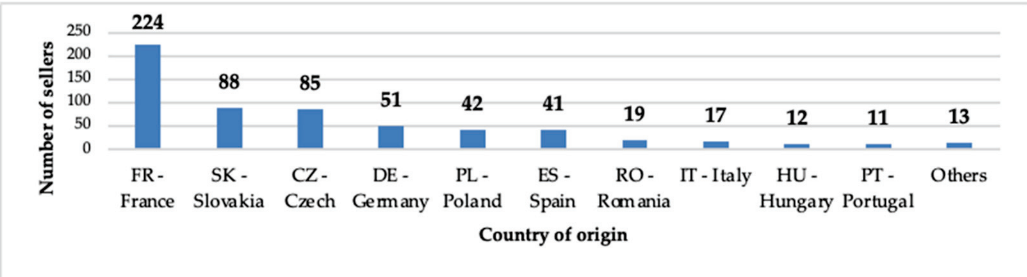

Figure 5. Supply chain for Citroen C3 and Peugeot 208 (2019). Source: Authors' compilation.

The outcomes of implementing the PSA TT (Trnava) supply chain in its entirety across Europe for the C3 and 208 models were displayed in 2019. There are 224 French companies (37.15%) and 85 Czech companies (14.1%). With 88 sellers, Slovak merchants are in the lead. Slovakia was PSA Group SVK TT's second-largest supplier of parts and components for automobiles (the C3 and 208) in 2019 (14.6%). There are now 603 sellers overall [67].

Figure 6 represents all PSA suppliers for 2020 and shows the total number of vendors by country of origin for three groups (particular suppliers for the C3, 208, and common suppliers for both models). France comes in first with 191 suppliers, or almost 31% of all suppliers, followed by Slovakia with 58 sellers (or about 9.43%), with Germany third (55, approx. 9%).

A global supply network exists here. Dealers (suppliers) have climbed to 615. Due to the fact that the cars were created and designed in France, this economy contributes the most to manufacturing and imports in both years while also producing the most added value overall. Slovakia manufactures, assembles, and transports these vehicles to the international market in addition to making important contributions to their manufacturing. It is the second-largest provider of these models in the world.

It is more effective to start from the fourth row, which is controlled by France because it imports connectors, with a volume of 148,793 pieces, while the first three firms only import helixes from the greatest suppliers by volume of imported pieces. With Adhex Technologies, Slovakia holds seventh spot and imports 104,253 foam parts. Plastic Omnium Auto Exteriors, which imports 1033 pieces of plastic parts, is the largest provider of different parts in France. Slovakia is in fifth position with the car seat and exhaust system importer Faurecia Automotive Slovakia s.r.o.

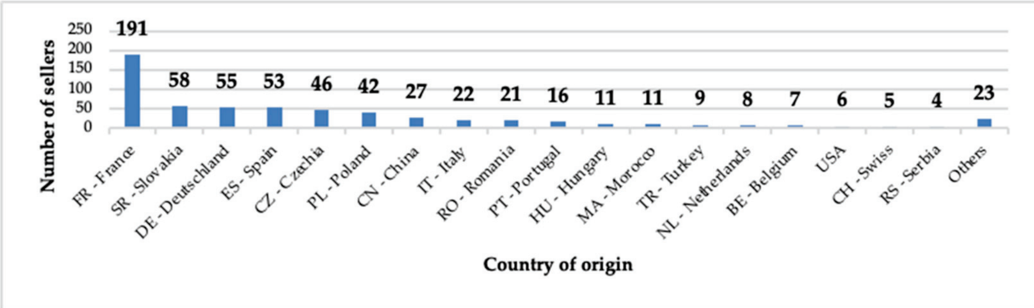

Figure 6. Supply chain for Citroen C3 and Peugeot 208 (2020). Source: Authors' compilation.

Slovakia does not import technologically and innovatively intensive materials and parts (Figure 1), instead concentrating on plastic production. The nations with the most sellers from the top 30 suppliers are included in the table below, along with the number of pieces supplied and part types. Tables 1 and 2 below show import comparisons [67].

Table 1. The largest suppliers of PSA Group SVK for 2019 in terms of volume of imported units and types of parts.

Imports Based on Volume		Imports Based on Types of Parts	
Country (Number of Suppliers)	**Volume (pcs)**	**Country (Number of Suppliers)**	**Types of Parts (pcs)**
FR (15)	5,805,554	FR (10)	2668
CZ (3)	364,377	CZ (6)	905
DE (1)	203,952	DE (1)	80
ES (1)	64,990	ES (1)	194
HU (1)	20,441	GB (1)	319
IT (1)	29,363	IT (1)	98
PL (3)	191,273	PL (3)	325
SK (5)	201,026	SK (7)	902

Source: Authors' compilation.

Table 2. The most important PSA Group SVK suppliers in 2019 in terms of the number of types of imported parts.

Order	Seller	Town	Products	Parts (pcs)
1.	Faurecia Automotive SVK s.r.o.	Trnava	Car seats	279
2.	Adhex Technologies	Senec	Foam parts	158
3.	Lear Corporation Seating SVK	Presov	Seating systems	119
4.	Eurostyle Systems s.r.o.	Banovce nad Bebravou	Plastic parts	95
5.	SMRC Automotive Solutions	Nitra	Modules, cockpits	92

Source: Authors' compilation.

France dominates in terms of the quantity of providers in both situations, importing 2668 different types of parts and totaling 5,805,554 pieces of material. The Czech Republic, in second position, imports 905 components and 364,377 volume units annually. Germany leads in terms of the quantity of imported goods, followed by third-placed Slovakia, which dominates when it comes to portion types. The final section's content is a list of the top 10 suppliers for 2019 in terms of the types of imported parts.

Faurecia Automotive SVK s.r.o. produces mainly automobile seats and exhaust systems, and imports 279 different types of parts, Adhex Technologies imports 158 foam parts, while Lear Corporation Seating Slovakia imports 119 parts for seating systems. It is a varied supply chain from Slovakia in both instances, with the majority being direct suppliers. With 58 suppliers, Slovakia contributed approximately 9.43% of the materials and components used in the production of automobiles in 2020. Plastic parts and components were the main imports. Low costs and a limited focus on science and research in the Slovak automotive industry are the key causes of the low quantity of imported technologically and innovation-intensive components. Applying an appropriate and comprehensive innovative strategy, particularly in regards higher quality education, is required to foster an increased first-rate productivity in Slovakia [67].

4.2. Implementing Industry 4.0 in PSA Group SVK

This part was developed with input from Industry 4.0 companies that digitalized their manufacturing processes (four enterprises, about 80 individuals). The respondents affirmed the relevance of cyber-physical production systems for almost all Slovak industries as well as their expectations for the sector's growth at the national level, particularly when analyzed from a long-term perspective. Implementing this idea and addressing it at the national level are crucial since innovation and investment in R&D may enhance the Slovak economy and industry [68]. The idea of cyber-physical production systems in relation to collision avoidance technologies, environment mapping algorithms, and mobility simulation tools in network connectivity systems is important for the Slovak industry since the automotive sector dominates the national economy and harnesses vehicle navigation and remote sensing technologies that make it possible for connected devices to communicate more quickly and effectively. The Internet of Things sensing networks are able to coordinate and receive multiple requests, data, and orders in real-time thanks to cloud and big data applications. Digitalization reduces the inefficient use of paper and other consumables while accelerating communication. As a result, there are fewer product flaws and there is greater control over the production process, by leveraging deep learning-assisted smart process planning, Internet of Things-based decision support systems, and cyber-physical system-based real-time monitoring [67].

Regarding the second query, an organization must comprehend the necessity of producing better and higher-quality products in Industry 4.0 more quickly and with fewer product faults. In the case of PSA Group SVK, which collaborates with Stellantis, the fourth-largest car manufacturer in the world, innovation is essential for the future. PSA Group SVK's capability for innovation will be further increased by this collaboration and growing capital expenditures. PSA Group SVK dominates the InoLab segment, with its primary goals including the configuration of manufacturing and logistics systems based on automation, digitalizing businesses, fostering collaboration with universities, tech firms, and governmental agencies, and managing EU funds, while collaborating with college students from Slovak and French institutions.

The most important element of this transition is the education and training of workers and students. The business management of PSA Group SVK claims that digitization produces a digital supply chain by more effectively, quickly, and efficiently networking goods, suppliers, manufacturers, and consumers. Digitalization is leveraged to improve employee communication with equipment and on the assembly line. Not everything has been digitalized yet, despite the time and work expended, but big improvements will soon be feasible. Several respondents noted the importance of new hire and continuous staff training, in addition to retraining.

The Industry 4.0 department of PSA Group SVK views cyber-physical production systems and artificial intelligence-based decision-making algorithms as designed to ensure that manufacturing processes communicate and exchange operational data, leading to unit autonomy and task optimization [69]. PSA Group SVK experimental projects are being developed by InoLab, which deploys robotic wireless sensor networks, product decision-making information systems, and Internet of Things smart devices. One illustration is the use of virtual reality to build a car and all of its parts. PSA Group SVK invests significantly in its workforce in terms of retraining and education, while making significant adjustments in supply structures. Although certainly not at all levels, PSA Group SVK became a digital firm by insisting that an idea be adopted by everyone involved in the whole production process. Slovakia has some of the most advanced environmental policies, complying with ISO 14,001 [67], the environmental management standard. PSA Group SVK follows stringent guidelines for the storage of chemical products as well as the discharge of wastewater or emissions into the atmosphere [70,71]. PSA Group SVK has fully digitalized its operations, while focusing mostly on the entire supply chain and digital interactions with suppliers. The following discusses the creation of technical product information, where it is useful to produce documents digitally and manage customer relationships [72].

The basic functions of assembly lines inside enterprises, such as the manufacture of vehicles, are the main areas of attention for automation. PSA Group SVK was able to implement groundbreaking innovations that have improved the efficiency of its operations with the introduction of the new generation of PSA automobiles, such as the Peugeot 208 (full kitting, laser geometry control, 675 robotic arms, edge supplying). These automation components are the main areas of concentration for PSA Group SVK, harnessing industrial big data, deep learning-assisted smart process planning, and real-time advanced analytics.

Figure 7 provides evidence that around 55% of the industrial sector is automated. The primary problems in this context are robotic processing and laser technology (675 robots). Maintenance and services come in jointly third with a 10% stake, while logistics is second with 20%. The other 5% of the activities cover various tasks, although the bulk (95%) of them are automated core solutions [67,73].

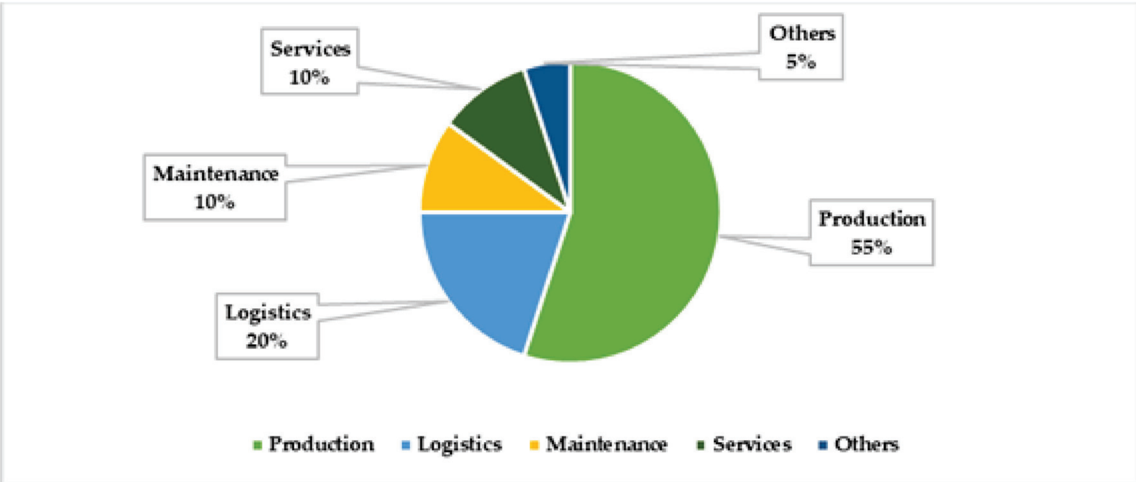

Figure 7. Areas of automation for PSA Group SVK. Source: Authors' compilation.

The COVID-19 pandemic has forced PSA Group SVK to cease all of its production. Given the continuous chip shortage and the present energy crisis, a shorter innovation cycle is also debatable. Once it has received the required information and an adequate supply of raw materials, PSA Group SVK can employ strategies to quicken the innovation cycle of the production process. Examples are statistical- and data-driven processes and production environment analyses [74]. Vehicles must be registered as they leave the final manufacturing line and head to the warehouse, where they will be transported in order as finished items to be recorded at PSA Group SVK. At that moment, an assigned employee creates a tangible record while simultaneously marking the completed model with a reader device (electric form and communication totem).

The corporate initiative Boost School AUT/ROB was launched in 2020 with the goal of enhancing the industrial automation and robotics competencies of maintenance personnel in manufacturing facilities. The program now has 49 PSA Group SVK students participating in dual education programs [67]. In 2021, 33,334 of all produced cars displayed the e-208 monogram and were electric. Electromobility currently drives every industry. The start-up costs for the Peugeot 208 and e-208 of the next generation were EUR 100 million. Trnava also saw the establishment of the first battery assembly facility. One of PSA Group SVK's most recent Industry 4.0 developments was the launch of InoLab in 2020, which accelerated the transition to paperless production and a digital supply chain. Furthermore, expenditures were made on comprehensive kitting, laser welding, and a structure for making batteries. Additionally noteworthy are PSA Group SVK's smooth press shops, environmentally friendly paint shops, and predictive maintenance technology.

For Industry 4.0 mobile maintenance management, PROCE55, an agile, state-of-the-art piece of software, was developed. System integration, reliable and unbiased machine data, and an online production overview are all provided, with a strong flexibility toward certain inventive strategies being dominant. By speeding production, improving quality, and facilitating post-production product control, Industry 4.0 technologies have the potential to greatly simplify the manufacturing process. PSA Group SVK leverages a plant quality indicator (DVT) to pinpoint problems [67,75]. The process will quicken and become more significant as more individuals express interest in electric automobiles. The Sector 4.0 team is in agreement that cutting-edge sustainable intelligent transportation systems, computer vision algorithms, and deep learning-based sensor technologies will shape the automobile industry in Slovakia.

In 2021, a new sector B production program at PSA Group SVK's Trnava manufacturing unit obtained financing. A new production plan for sector B will gradually begin in 2023. A substantial portion of the production program will be made up entirely of electric motors to greatly increase carbon neutrality. Industry 4.0 technological application, innovation, energy intensity reduction, and environmental protection will significantly grow as a result of the industrial investment in the new production program [76]. EU funds for employee training have been beneficial for PSA Group SVK. By integrating immersive digital simulations, Industry 4.0′s benefits in terms of visual and spatial analytics will include growing competition, decreased costs and inventories, and increased production efficiency through cognitive artificial intelligence algorithms [77]. One drawback highlighted by respondents was the probable loss of some work prospects. They also affirmed the urgency of putting this strategy into action, particularly in light of the growing rivalry between nearby economies.

Business data for the years 2020–2021 analysis showed that the Peugeot 208 and Citroen C3 vehicles have contributed value that spans several economies. France generated the largest value (31.06%), followed by Slovakia (around 9.5%) (primarily assembly labor; see Figure 8). In France, vehicle engineering and design employ the most recent technology. The contribution of Slovak value to the industrial process is likely to rise annually.

We used InoLab to create a projection of prospective growth in order to compare the amount of added value before and after the implementation of Industry 4.0-based manufacturing systems at the Trnava facility, leveraging cyber-physical production systems, spatial data visualization and cognitive data fusion techniques, and virtual reality modeling tools.

Industry 4.0 as a whole and its implications are necessary to achieve higher added value and harnessing cyber-physical production systems can also lead to the desired outcomes for other sectors. After the integration of the Internet of Things-based decision support systems across Slovak plants, the share of national added value significantly increased to almost 20% in the past 2 years. Other economies, particularly in the EU, will see similar outcomes to this study, given that most of them place a strong emphasis on the automobile sector (Figure 9).

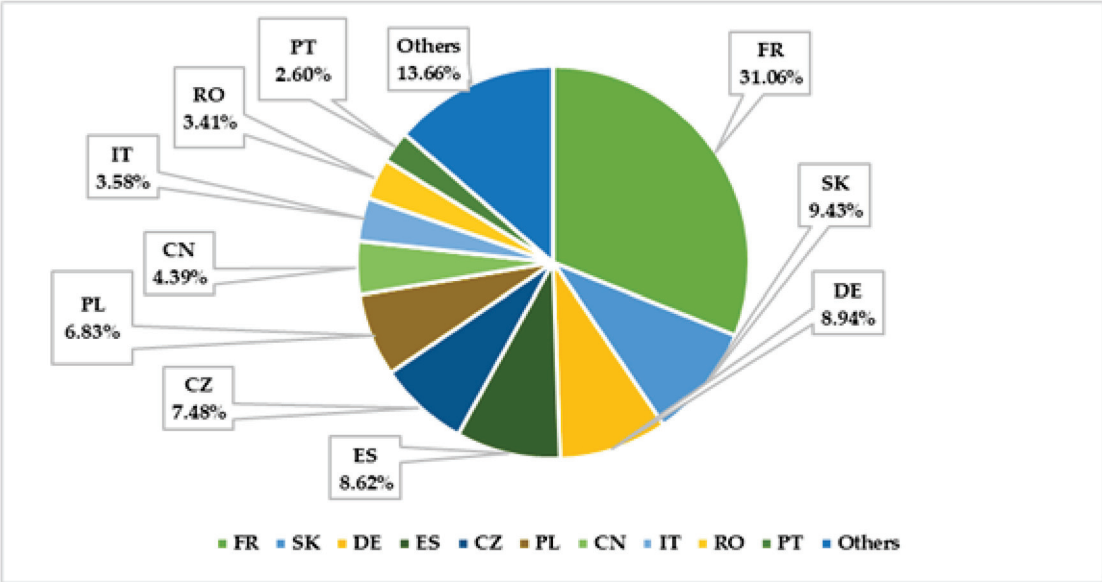

Figure 8. Value-added percentages (2020–2021) for Citroen C3 and Peugeot 208 automobiles, broken down by nation before implementing I4 trends. Source: Authors' compilation.

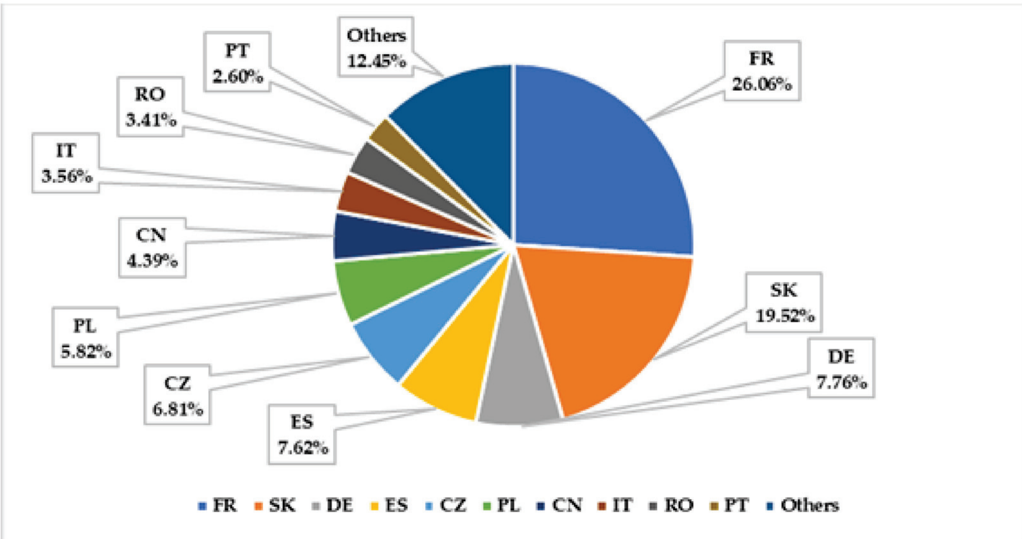

Figure 9. Value-added percentages (2021–2023) for C3 and 208 automobiles, broken down by nation after implementing Industry 4.0 trends. Source: Authors' compilation.

By deploying geospatial data mining and virtual simulation modeling tools, sensing and computing technologies, and computer vision algorithms, Slovak automotive companies will raise the added value of their products and make them more marketable [78]. Automakers will become inventive as a result and can become the industry leaders. Uncertainty exists as to whether and how the Slovak government and general legislation will be able to address this trend of supporting businesses with creative policies, improved regulations, and specific norms. Respondents view this as a weakness of the Slovak Industry 4.0-based manufacturing systems. Once businesses are able to finance their own R&D, Slovakia's added value will grow rapidly [79].

Industry 4.0-based manufacturing systems have to be broadly adopted in order for all industries to both comprehend and gain from it. The goal of the Slovak automotive industry is to combine activities for R&D with innovation, by leveraging cognitive wireless sensor networks, connected vehicle technologies, environment mapping and location tracking algorithms, and mapping and navigation tools in smart urban mobility systems. The quality of life in Slovakia will be improved as a result of cyber-physical production systems, with enterprises able to integrate predictive maintenance systems, robotic wireless sensor networks, and cognitive automation. To configure Industry 4.0-based manufacturing systems developed on machine and deep learning technologies, spatial simulation and motion planning algorithms, and traffic flow prediction tools, in-depth studies must be performed, and a Slovak Intelligent Industry Platform should be set up [80] (Table 3).

Table 3. A thorough list of suggestions for the Slovak government.

Areas	Recommendations
Increasing understanding and cooperating	1. Campaign for information as regards cyber-physical production systems 2. Encouragement of Internet of Things-based real-time production logistics research 3. Industry 4.0-based manufacturing systems implementation guidebook 4. Increased promotion of sustainable Internet of Manufacturing Things
Industry 4.0 Research	1. Assistance with applied research 2. Research agenda for Industry 4.0 3. Sector-oriented consortia 4. Attempts to cut back on R&D expenses
The Smart Factory	1. Support for the use of innovative materials and technologies 2. Standardization (reference architecture) 3. Introduction of new models into supply chains 4. Use of industrial big data analytics
Financing	1. Better funding mechanisms 2. Address the needs of the research agenda 3. Innovative public procurement 4. Implementation of pilot projects
Employment and education	1. A breakdown of the current situation’s primary needs 2. Creating predictive curricula 3. Providing more specialized skills 4. Following the European agenda as regards new skills
E-government and legislation	1. Talent development that is ongoing in the public sector 2. Commercial big data usage (Big Data) 3. Government’s active involvement in promoting the implementation of Industry 4.0 4. A suggestion for a clear vs. digitization strategy

Source: Authors’ compilation of collected data analysis.

An action plan tailored for a particular location would be the key document of this platform [80]. This strategy would articulate the platform and establish long-term objectives in the fields of multiple energy, materials, nanotechnology, and robotics techniques, while addressing environmental policy, which is equally crucial for progress [81].

Suggestions for environmental politics:

1. In order to draw in international investment, the Slovak government must provide favorable circumstances for enterprises to go green.
2. Slovakian business must exert pressure on the government to establish the necessary legal framework for environmental protection.
3. The vehicle manufacturers’ headquarters must collaborate with their Slovakian suppliers to assist them in retraining workers to take advantage of new technology and production methods [82].
4. New training programs and cross-sectoral collaboration between the public and private sectors as well as academia are required for retraining and enhancing the quality of personnel to fulfil the work requirements of the rising e-mobility sub-sectors [83].

Cyber-physical smart manufacturing systems, knowledge acquisition-based organizational achievements, data visualization tools, and sustainable economic development configure Industry 4.0 wireless networks across immersive work environments in the Slovak automotive sector [84–87]. Artificial intelligence data-driven Internet of Things systems, tradeable digital assets, and decision intelligence and modeling articulate sustainable smart manufacturing. Real-time advanced analytics, immersive extended reality technologies, and socially interconnected virtual services assist cyber-physical production networks [88–91]. Geospatial big data management algorithms, decision intelligence and modeling, and blockchain technology adoption enable deep learning-assisted smart process planning. Robotic wireless sensor networks and sensory algorithmic devices further Internet of Manufacturing Things [92–95]. Immersive virtual technologies, virtual marketplace dynamics data, knowledge co-creation, and remote working tools further spatial analytics [96–99]. Sensing technologies, international business performance, knowledge capitalism, and cognitive analytics management shape cyber-physical manufacturing and immersive visualization systems in the Slovak automotive sector [100–103].

Industry 4.0-based car and part manufacturing systems deploy artificial intelligence-based decision-making algorithms, visual perception and remote sensing technologies, route detection and object localization technologies, digital twin-based product development, real-world connected vehicle navigation data, big geospatial data analytics, and trajectory planning and mobility simulation tools throughout the automotive sector and networked transport systems [104–107]. The Internet of Things sensing networks, spatial data visualization tools, cyber-physical production networks, and cognitive decision-making algorithms configure the Industry 4.0-based automotive sector. Urban transportation systems integrate Internet of Things-based real-time production logistics, virtual reality modeling tools, machine and deep learning technologies, and digital twin algorithms [108–111]. Spatial data visualization techniques, robotic wireless sensor networks, geospatial data mining and virtual simulation modeling tools, and cognitive artificial intelligence algorithms optimize smart sustainable urban mobility systems [112–115].

5. Conclusions

Our research findings, and the novelty brought to the topic of the Industry 4.0 manufacturing systems, are related to the integration of computer vision algorithms, remote sensing data fusion techniques, and mapping and navigation tools in the Slovak automotive sector. We conducted a thorough examination of Industry 4.0-based value and supply chains, clarifying how cyber-physical production systems operate in relation to collision avoidance technologies, environment mapping algorithms, and mobility simulation tools in network connectivity systems.

Based on the comparative analysis, Slovakia, which is second to France in terms of added value, has the biggest number of suppliers. In order to boost the nation’s total industry and offer even more added value, both Industry 4.0 and R&D on a large scale should be addressed. Data visualization tools, cognitive data fusion techniques, and deep learning-based computer vision algorithms are pivotal in digital twin-based cyber-physical production systems in the Slovak automotive sector. Cloud computing and remote sensing technologies, immersive visualization tools, and cyber-physical production systems can be operational in virtual enterprises and immersive 3D environments by use of virtual modeling and simulation tools, digital twin modeling, and neural network algorithms in the Slovak automotive sector. Industry 4.0 has a substantial impact on automotive companies’ product range expansion with value-added growth and highlights a lack of governmental support, particularly in the areas of legislation, financing, R&D, and education. Industry 4.0-based manufacturing systems can boost Slovakia’s need for innovation, deployment of cutting-edge technologies, and programmatic changes in training in order to achieve high added value in the automotive sector and subsequent exports of goods and services through a consistent action plan. Slovakia can develop the proactive nature of an innovative nation based on putting these suggestions and recommendations into practice; otherwise, it will continue to be an assembly country.

6. Specific Contributions to the Literature

Mapping and navigation tools, machine and deep learning technologies, artificial intelligence data-driven Internet of Things systems, and cognitive wireless sensor networks articulate the smart sustainable urban transport architecture. The Industry 4.0-based automotive sector integrates connected vehicle technologies, intelligent transportation planning tools, the Internet of Things sensing networks, and spatial simulation and motion planning algorithms. Traffic flow prediction tools, deep learning-assisted smart process planning, industrial big data, and the Internet of Things-based real-time production logistics are pivotal in the Industry 4.0-based Slovak automotive sector.

7. Limitations and Further Directions of Research

This study has some limitations, mainly because the analysis covered a single country (Slovakia) and a single company (PSA Group SVK). Moreover, more investigations are required to articulate Industry 4.0-based manufacturing systems across the Slovak automotive sector, taking into account self-driving vehicles in relation to smart transportation and network connectivity systems, in terms of deep learning object detection and collision avoidance technologies, geospatial data visualization tools, and trajectory planning and sensor fusion algorithms.

8. Practical Implications

Autonomous manufacturing processes, interconnected virtual services, traffic flow prediction tools, and collision avoidance technologies shape intelligent transportation planning and engineering. Remote sensing technologies, obstacle avoidance algorithms, and mapping and navigation tools build on artificial intelligence-based decision-making algorithms, cyber-physical system-based real-time monitoring tools, and cognitive automation throughout the automotive sector and networked transport. Artificial intelligence-based decision-making algorithms, connected vehicle technologies, sensing and computing tools, and cyber-physical production systems are instrumental in the Industry 4.0-based Slovak automotive sector.

Author Contributions: Conceptualization, M.N. and G.L.; methodology, M.N. and G.L.; validation, M.N. and G.L.; formal analysis, M.N.; investigation, M.N.; resources, M.N. and G.L.; data curation, G.L.; writing—original draft preparation, M.N. and G.L.; writing—review and editing, M.N. and G.L.; visualization, M.N.; supervision, G.L. All authors have read and agreed to the published version of the manuscript.

Funding: This paper is an output of the project NFP313010BWN6 "The implementation framework and business model of the Internet of Things, Industry 4.0 and smart transport".

Data Availability Statement: The data presented in this study are available on request from the corresponding author.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Hatzigeorgiou, A.; Lodefalk, M. A literature review of the nexus between migration and internationalization. *J. Int. Trade Econ. Dev.* **2021**, *30*, 319–340. [CrossRef]
2. Nikulin, D.; Wolszczak-Derlacz, J.; Parteka, A. GVC and wage dispersion. Firm level evidence from employee-employer database. *Equilib. Q. J. Econ. Econ. Policy* **2021**, *16*, 357–375. [CrossRef]
3. Kordalska, A.; Olczyk, M. New patterns in the position of CEE countries in global value chains: Functional specialisation approach. *Oeconomia Copernic.* **2021**, *12*, 35–52. [CrossRef]
4. Minarik, M.; Zabojnik, S.; Pasztorova, J. Sources of Value-Added in V4 automotive GVCs: The Case of Transport and Storage Services and Firm Level Technology Absorption. *Cent. Eur. Bus. Rev.* **2022**, *11*, 12–14. [CrossRef]
5. Schoeneman, J.; Zhou, B.L.; Desmarais, B.A. Complex dependence in foreign direct investment: Network theory and empirical analysis. *Political Sci. Res. Methods* **2022**, *12*, 243–259. [CrossRef]
6. Kumar, S.; Raut, R.D.; Narwane, V.S.; Narkhede, B.E. Applications of industry 4.0 to overcome the COVID-19 operational challenges. *Diabetes Metab. Syndr. Clin. Res. Rev.* **2020**, *14*, 1283–1289. [CrossRef]

7. Said, M.; Shaheen, A.M.; Ginidi, A.R.; El-Sehiemy, R.A.; Mahmoud, K.; Lehtonen, M.; Darwish, M.M.F. Estimating Parameters of Photovoltaic Models Using Accurate Turbulent Flow of Water Optimizer. *Processes* **2021**, *9*, 627. [CrossRef]
8. Martinez-Noya, A.; Garcia-Canal, E. International evidence on R&D services outsourcing practices by technological firms. *Multinatl. Bus. Rev.* **2014**, *22*, 372–393.
9. Bonab, A.-F. The Development of Competitive Advantages of Brand in the Automotive Industry (Case Study: Pars Khodro Co). *J. Internet Bank. Commer.* **2017**, *22*, S8.
10. Pavlinek, P.; Zenka, J. Value creation and value capture in the automotive industry: Empirical evidence from Czechia. *Environ. Plan.* **2016**, *48*, 937–959. [CrossRef]
11. Modibbo, U.M.; Gupta, N.; Chatterjee, P.; Ali, I. A Systematic Review on the Emergence and Applications of Industry 4.0. In *Computational Modelling in Industry 4.0.*; Springer: Singapore, 2022. [CrossRef]
12. Stock, T.; Obenaus, M.; Kunz, S.; Kohl, H. Industry 4.0 as enabler for a sustainable development: A qualitative assessment of its ecological and social potential. *Process Saf. Environ. Prot.* **2018**, *118*, 254–267. [CrossRef]
13. Kotlebova, J.; Arendas, P.; Chovancova, B. Government expenditures in the support of technological innovations and impact on stock market and real economy: The empirical evidence from the US and Germany. *Equilib. Q. J. Econ. Econ. Policy* **2020**, *15*, 717–734. [CrossRef]
14. Cerna, I.; Elteto, A.; Folfas, P.; Kuznar, A.; Krenkova, E.; Minarik, M.; Przezdziecka, E.; Szalavetz, A.; Tury, G.; Zabojnik, S. *GVCs in Central Europe—A Perspective of the Automotive Sector after COVID-19*; Ekonom: Bratislava, Slovakia, 2022; Available online: https://gvcsv4.euba.sk/images/PDF/monograph.pdf (accessed on 4 June 2022).
15. Gereffi, G.; Humphrey, J.; Sturgeon, T. The governance of global value chains. *Rev. Int. Political Econ.* **2005**, *12*, 78–104. [CrossRef]
16. Pugliese, E.; Napolitano, L.; Zaccaria, A.; Pietronero, L. Coherent diversification in corporate technological portfolios. *PLoS ONE* **2019**, *14*, e0223403. [CrossRef]
17. Mondejar, M.A.; Avtar, R.; Diaz, H.L.; Dubey, R.K.; Esteban, J.; Gomez-Morales, A.; Hallam, B.; Mbungu, N.T.; Okolo, C.C.; Prasad, K.A.; et al. Digitalization to achieve sustainable development goals: Steps towards a Smart Green Planet. *Sci. Total Environ.* **2021**, *794*, 148539. [CrossRef]
18. Mugge, D. International economic statistics: Biased arbiters in global affairs? *Fudan J. Humanit. Soc. Sci.* **2020**, *13*, 93–112. [CrossRef]
19. Ye, C.S.; Ye, Q.; Shi, X.P.; Sun, Y.P. Technology gap, global value chain and carbon intensity: Evidence from global manufacturing industries. *Energy Policy* **2020**, *137*, 111094. [CrossRef]
20. Brioschi, M.; Bonardi, M.; Fabrizio, N.; Fuggetta, A.; Vrga, E.S.; Zuccala, M. Enabling and Promoting Sustainability through Digital API Ecosystems: An example of successful implementation in the smart city domain. *Technol. Innov. Manag. Rev.* **2021**, *11*, 4–10. [CrossRef]
21. Kubickova, L.; Kormanakova, M.; Vesela, L.; Jelinkova, Z. The Implementation of Industry 4.0 Elements as a Tool Stimulating the Competitiveness of Engineering Enterprises. *J. Compet.* **2021**, *13*, 76–94. [CrossRef]
22. Tao, F.; Qi, Q.; Wan, L.; Nee, A.Y.C. Digital Twins and Cyber-Physical Systems toward Smart Manufacturing and Industry 4.0: Correlation and Comparison. *Engineering* **2019**, *5*, 653–661. [CrossRef]
23. Sjodin, D.R.; Parida, V.; Leksell, M.; Petrovic, A. Smart Factory Implementation and Process Innovation. *Res. Technol. Manag.* **2018**, *61*, 22–31. [CrossRef]
24. Gray, M.; Kovacova, M. Internet of Things Sensors and Digital Urban Governance in Data-driven Smart Sustainable Cities. *Geopolit. Hist. Int. Relat.* **2021**, *13*, 107–120.
25. Kliestik, T.; Zvarikova, K.; Lazaroiu, G. Data-driven Machine Learning and Neural Network Algorithms in the Retailing Environment: Consumer Engagement, Experience, and Purchase Behaviors. *Econ. Manag. Financ. Mark.* **2022**, *17*, 57–69. [CrossRef]
26. Andronie, M.; Lazaroiu, G.; Iatagan, M.; Uța, C.; Ștefanescu, R.; Cocoșatu, M. Artificial Intelligence-Based Decision-Making Algorithms, Internet of Things Sensing Networks, and Deep Learning-Assisted Smart Process Management in Cyber-Physical Production Systems. *Electronics* **2021**, *10*, 2497. [CrossRef]
27. Skare, M.; Gil-Alana, L.A.; Claudio-Quiroga, G.; Prziklas Druzeta, R. Income inequality in China 1952-2017: Persistence and main determinants. *Oeconomia Copernic.* **2021**, *12*, 863–888. [CrossRef]
28. Popescu, G.H.; Valaskova, K.; Horak, J. Augmented Reality Shopping Experiences, Retail Business Analytics, and Machine Vision Algorithms in the Virtual Economy of the Metaverse. *J. Self Gov. Manag. Econ.* **2022**, *10*, 67–81. [CrossRef]
29. Zabojnik, S. *Selected Problems of International Trade and International Business*; Econom: Bratislava, Slovakia, 2015; pp. 28–61.
30. Krykavskyy, A.; Pokhylchenko, O.; Hayvanovych, N. Supply chain development drivers in industry 4.0 in Ukrainian enterprises. *Oeconomia Copernic.* **2019**, *10*, 273–290. [CrossRef]
31. Clayton, E.; Kral, P. Autonomous Driving Algorithms and Behaviors, Sensing and Computing Technologies, and Connected Vehicle Data in Smart Transportation Networks. *Contemp. Read. Law Soc. Justice* **2021**, *13*, 9–22.
32. Yang, F.; Gu, S. Industry 4.0, a revolution that requires technology and national strategies. *Complex Intell. Syst.* **2021**, *7*, 1311–1325. [CrossRef]
33. Hermann, M.; Pentek, T.; Otto, B. Design Principles for Industry 4.0 Scenarios: A Literature Review. Technische Universitat Dortmund. 2016. Available online: http://www.snom.mb.tu-dortmund.de/cms/de/forschung/Arbeitsberichte/Design-Principles-for-Industrie-4_0-Scenarios.pdf (accessed on 2 May 2022).

34. Klingenberg, C.O.; Borges, M.A.V.; Antunes, J., Jr. Industry 4.0 as a data-driven paradigm: A systematic literature review on technologies. *J. Manuf. Technol. Manag.* **2019**, *32*, 570–592. [CrossRef]
35. Rogers, S.; Zvarikova, K. Big Data-driven Algorithmic Governance in Sustainable Smart Manufacturing: Robotic Process and Cognitive Automation Technologies. *Anal. Metaphys.* **2021**, *20*, 130–144.
36. Ruttimann, B.G.; Stockli, M.T. Lean and Industry 4.0—Twins, Partners, or Contenders? A Due Clarification Regarding the Supposed Clash of Two Production Systems. *J. Sci. Serv. Manag.* **2016**, *9*, 485–500. [CrossRef]
37. Krulicky, T.; Horak, J. Business performance and financial health assessment through artificial intelligence. *Ekon. Manaz. Spektrum* **2021**, *15*, 38–51. [CrossRef]
38. Kovacova, M.; Lewis, E. Smart Factory Performance, Cognitive Automation, and Industrial Big Data Analytics in Sustainable Manufacturing Internet of Things. *J. Self Gov. Manag. Econ.* **2021**, *9*, 9–21.
39. Mehmann, J.; Teuteberg, F. The fourth-party logistics service provider approach to support sustainable development goals in transportation—A case study of the German agricultural bulk logistics sector. *J. Clean. Prod.* **2016**, *126*, 382–393. [CrossRef]
40. Galbraith, A.; Podhorska, I. Artificial Intelligence Data-driven Internet of Things Systems, Robotic Wireless Sensor Networks, and Sustainable Organizational Performance in Cyber-Physical Smart Manufacturing. *Econ. Manag. Financ. Mark.* **2021**, *16*, 56–69.
41. Sony, M. Pros and cons of implementing Industry 4.0 for the organizations: A review and synthesis of evidence. *Prod. Manuf. Res.* **2020**, *8*, 244–272. [CrossRef]
42. Dalenogare, L.; Benitez, N.; Ayala, N.F.; Frank, A.G. The expected contribution of Industry 4.0 technologies for industrial performance. *Int. J. Prod. Econ.* **2018**, *204*, 383–394. [CrossRef]
43. Hinds, P.S.; Vogel, R.J.; Clarke-Steffen, L. The possibilities and pitfalls of doing a secondary analysis of a qualitative data set. *Qual. Health Res.* **1997**, *7*, 408–424. [CrossRef]
44. Svabova, L.; Tesarova, E.N.; Durica, M.; Strakova, L. Evaluation of the impacts of the COVID-19 pandemic on the development of the unemployment rate in Slovakia: Counterfactual before-after comparison. *Equilib. Q. J. Econ. Econ. Policy* **2021**, *16*, 261–284. [CrossRef]
45. Verhof, P.C.; Broekhuizen, T.; Bart, Y.; Bhattacharya, A.; Dong, J.Q.; Fabian, N.; Haenlein, M. Digital transformation: A multidisciplinary reflection and research agenda. *J. Bus. Res.* **2021**, *122*, 889–901. [CrossRef]
46. Lawrence, J.; Durana, P. Artificial Intelligence-driven Big Data Analytics, Predictive Maintenance Systems, and Internet of Things-Based Real-Time Production Logistics in Sustainable Industry 4.0 Wireless Networks. *J. Self Gov. Manag. Econ.* **2021**, *9*, 62–75.
47. Zavadska, Z.; Zavadsky, J. *Industry 4.0 and Intelligent Technologies in the Development of the Corporate Operation Management*; Belianum: Banska Bystrica, Slovakia, 2020; pp. 130–155.
48. Schot, J.; Steinmueller, W.E. Three frames for innovation policy: R&D, systems of innovation and transformative change. *Res. Policy* **2018**, *47*, 1554–1567.
49. Lazaroiu, G.; Harrison, A. Internet of Things Sensing Infrastructures and Data-driven Planning Technologies in Smart Sustainable City Governance and Management. *Geopolit. Hist. Int. Relat.* **2021**, *13*, 23–36.
50. Vinerean, S.; Budac, C.; Baltador, L.A.; Dabija, D.-C. Assessing the Effects of the COVID-19 Pandemic on M-Commerce Adoption: An Adapted UTAUT2 Approach. *Electronics* **2022**, *11*, 1269. [CrossRef]
51. Zhong, R.; Xu, X.; Klotz, E.; Newman, S.T. Intelligent Manufacturing in the Context of Industry 4.0: A Review. *Engineering* **2021**, *3*, 616–630. [CrossRef]
52. Belhadi, A.; Kamble, S.; Jabbour, C.J.C.; Gunasekaran, A.; Ndubisi, N.O.; Venkatesh, M. Manufacturing and service supply chain resilience to the COVID-19 outbreak: Lessons learned from the automobile and airline industries. *Technol. Forecast. Soc. Change* **2021**, *163*, 120–447. [CrossRef]
53. Hoffmann, M. *Smart Agents for the Industry 4.0*; Springer: Berlin, Germany, 2019; pp. 226–297.
54. State of FDI in Slovakia. 2022. Available online: https://www.sario.sk/sk/investicie/pzi-pribehy-uspesnych/pzi-prilev-odlev (accessed on 28 May 2022).
55. Hopkins, E.; Siekelova, A. Internet of Things Sensing Networks, Smart Manufacturing Big Data, and Digitized Mass Production in Sustainable Industry 4.0. *Econ. Manag. Financ. Mark.* **2021**, *16*, 28–41.
56. Pisar, P.; Bilkova, D. Controlling as a tool for SME management with an emphasis on innovations in the context of Industry 4.0. *Equilib. Q. J. Econ. Econ. Policy* **2019**, *14*, 763–785. [CrossRef]
57. Johnson, E.; Nica, E. Connected Vehicle Technologies, Autonomous Driving Perception Algorithms, and Smart Sustainable Urban Mobility Behaviors in Networked Transport Systems. *Contemp. Read. Law Soc. Justice* **2021**, *13*, 37–50. [CrossRef]
58. Vuong, T.K.; Mansori, S. An Analysis of the Effects of the Fourth Industrial Revolution on Vietnamese Enterprises. *Manag. Dyn. Knowl. Econ.* **2021**, *9*, 447–459.
59. Chang, B.G.; Wu, K.S. The nonlinear relationship between financial flexibility and enter-prise risk-taking during the COVID-19 pandemic in Taiwan's semiconductor industry. *Oeconomia Copernic.* **2021**, *12*, 307–333. [CrossRef]
60. Lazaroiu, G.; Kliestik, T.; Novak, A. Internet of Things Smart Devices, Industrial Artificial Intelligence, and Real-Time Sensor Networks in Sustainable Cyber-Physical Production Systems. *J. Self Gov. Manag. Econ.* **2021**, *9*, 20–30. [CrossRef]
61. Sierra-Perez, J.; Teixeira, J.G.; Romero-Piqueras, C.; Patricio, L. Designing sustainable services with the ECO-Service design method: Bridging user experience with environmental performance. *J. Clean. Prod.* **2021**, *305*, 127228. [CrossRef]

62. European Commision, European Competitiveness Report 2014–2021. 2022. Available online: http://ec.europa.eu/enterprise/policies/industrial-competitiveness/competitiveness-analysis/european-competitiveness-report/index_en.htm (accessed on 29 April 2022).
63. Hamilton, S. Deep Learning Computer Vision Algorithms, Customer Engagement Tools, and Virtual Marketplace Dynamics Data in the Metaverse Economy. *J. Self Gov. Manag. Econ.* **2022**, *10*, 37–51. [CrossRef]
64. Konecny, V.; Barnett, C.; Poliak, M. Sensing and Computing Technologies, Intelligent Vehicular Networks, and Big Data-Driven Algorithmic Decision-Making in Smart Sustainable Urbanism. *Contemp. Read. Law Soc. Justice* **2021**, *13*, 30–39. [CrossRef]
65. Lyons, N. Deep Learning-based Computer Vision Algorithms, Immersive Analytics and Simulation Software, and Virtual Reality Modeling Tools in Digital Twin-Driven Smart Manufacturing. *Econ. Manag. Financ. Mark.* **2022**, *17*, 67–81. [CrossRef]
66. Glogovețan, A.I.; Dabija, D.-C.; Fiore, M.; Pocol, C.B. Consumer Perception and Understanding of European Union Quality Schemes: A Systematic Literature Review. *Sustainability* **2022**, *14*, 1667. [CrossRef]
67. PSA Slovakia. 2022. Available online: http://www.psa-slovakia.sk/o-nas.html?page_id=172 (accessed on 15 March 2022).
68. Nica, E. Urban Big Data Analytics and Sustainable Governance Networks in Integrated Smart City Planning and Management. *Geopolit. Hist. Int. Relat.* **2021**, *13*, 93–106. [CrossRef]
69. Poliak, M.; Poliakova, A.; Svabova, L.; Zhuravleva, A.N.; Nica, E. Competitiveness of Price in International Road Freight Transport. *J. Compet.* **2021**, *13*, 83–98. [CrossRef]
70. Kovacova, M.; Lazaroiu, G. Sustainable Organizational Performance, Cyber-Physical Production Networks, and Deep Learning-assisted Smart Process Planning in Industry 4.0-based Manufacturing Systems. *Econ. Manag. Financ. Mark.* **2021**, *16*, 41–54. [CrossRef]
71. Kovacova, M.; Machova, V.; Bennett, D. Immersive Extended Reality Technologies, Data Visualization Tools, and Customer Behavior Analytics in the Metaverse Commerce. *J. Self Gov. Manag. Econ.* **2022**, *10*, 7–21. [CrossRef]
72. Popescu, G.H.; Poliak, M.; Manole, C.; Dumitrescu, C.-O. Decentralized Finance, Blockchain Technology, and Digital Assets in Non-Fungible Token (NFT) Markets. *Smart Gov.* **2022**, *1*, 64–78. [CrossRef]
73. Wallace, S.; Lazaroiu, G. Predictive Control Algorithms, Real-World Connected Vehicle Data, and Smart Mobility Technologies in Intelligent Transportation Planning and Engineering. *Contemp. Read. Law Soc. Justice* **2021**, *13*, 79–92. [CrossRef]
74. Poliak, M.; Poliakova, A.; Zhuravleva, N.A.; Nica, E. Identifying the Impact of Parking Policy on Road Transport Economics. *Mob. Netw. Appl.* **2021**. *early access*. [CrossRef]
75. Kliestik, T.; Valaskova, K.; Nica, E.; Kovacova, M.; Lazaroiu, G. Advanced Methods of Earnings Management: Monotonic Trends and Change-Points under Spotlight in the Visegrad Countries. *Oeconomia Copernic.* **2020**, *11*, 371–400. [CrossRef]
76. Cooper, H.; Poliak, M.; Konecny, V. Computationally Networked Urbanism and Data-driven Planning Technologies in Smart and Environmentally Sustainable Cities. *Geopolit. Hist. Int. Relat.* **2021**, *13*, 20–30. [CrossRef]
77. Kliestik, T.; Novak, A.; Lazaroiu, G. Live Shopping in the Metaverse: Visual and Spatial Analytics, Cognitive Artificial Intelligence Techniques and Algorithms, and Immersive Digital Simulations. *Linguist. Philos. Investig.* **2022**, *21*, 187–202. [CrossRef]
78. Woodward, B.; Kliestik, T. Intelligent Transportation Applications, Autonomous Vehicle Perception Sensor Data, and Decision-Making Self-Driving Car Control Algorithms in Smart Sustainable Urban Mobility Systems. *Contemp. Read. Law Soc. Justice* **2021**, *13*, 51–64. [CrossRef]
79. Higgins, M.; Horak, J. Cyber-Physical Process Monitoring Systems, Artificial Intelligence-based Decision-Making Algorithms, and Sustainable Industrial Big Data in Smart Networked Factories. *Econ. Manag. Financ. Mark.* **2021**, *16*, 42–55. [CrossRef]
80. Franklin, K.; Potcovaru, A.M. Autonomous Vehicle Perception Sensor Data in Sustainable and Smart Urban Transport Systems. *Contemp. Read. Law Soc. Justice* **2021**, *13*, 101–110. [CrossRef]
81. Markauskas, M.; Baliute, A. Technological progress spillover effect in Lithuanian manufacturing industry. *Equilib. Q. J. Econ. Econ. Policy* **2021**, *16*, 783–806. [CrossRef]
82. Strakova, J.; Simberova, I.; Partlova, P.; Vachal, J.; Zich, R. The Value Chain as the Basis of Business Model Design. *J. Compet.* **2021**, *13*, 135–151. [CrossRef]
83. Gajdosikova, D.; Valaskova, K.; Kliestik, T.; Machova, V. COVID-19 Pandemic and Its Impact on Challenges in the Construction Sector: A Case Study of Slovak Enterprises. *Mathematics* **2022**, *10*, 3130. [CrossRef]
84. Valaskova, K.; Nagy, M.; Zabojnik, S.; Lazaroiu, G. Industry 4.0 Wireless Networks and Cyber-Physical Smart Manufacturing Systems as Accelerators of Value-Added Growth in Slovak Exports. *Mathematics* **2022**, *10*, 2452. [CrossRef]
85. Vatamanescu, E.-M.; Bratianu, C.; Dabija, D.-C.; Popa, S. Capitalizing Online Knowledge Networks: From Individual Knowledge Acquisition towards Organizational Achievements. *J. Knowl. Manag.* **2022**. [CrossRef]
86. Zvarikova, K.; Cug, J.; Hamilton, S. Virtual Human Resource Management in the Metaverse: Immersive Work Environments, Data Visualization Tools and Algorithms, and Behavioral Analytics. *Psychosociological Issues Hum. Resour. Manag.* **2022**, *10*, 7–20. [CrossRef]
87. Ionescu, L. Leveraging Green Finance for Low-Carbon Energy, Sustainable Economic Development, and Climate Change Mitigation during the COVID-19 Pandemic. *Rev. Contemp. Philos.* **2021**, *20*, 175–186. [CrossRef]
88. Watson, R. Tradeable Digital Assets, Immersive Extended Reality Technologies, and Blockchain-based Virtual Worlds in the Metaverse Economy. *Smart Gov.* **2022**, *1*, 7–20. [CrossRef]
89. Nica, E.; Poliak, M.; Popescu, G.H.; Parvu, I.-A. Decision Intelligence and Modeling, Multisensory Customer Experiences, and Socially Interconnected Virtual Services across the Metaverse Ecosystem. *Linguist. Philos. Investig.* **2022**, *21*, 137–153. [CrossRef]

90. Durana, P.; Perkins, N.; Valaskova, K. Artificial Intelligence Data-driven Internet of Things Systems, Real-Time Advanced Analytics, and Cyber-Physical Production Networks in Sustainable Smart Manufacturing. *Econ. Manag. Financ. Mark.* **2021**, *16*, 20–30.
91. Małkowska, A.; Urbaniec, M.; Kosała, M. The impact of digital transformation on European countries: Insights from a comparative analysis. *Equilib. Q. J. Econ. Econ. Policy* **2021**, *16*, 325–355. [CrossRef]
92. Lazaroiu, G.; Andronie, M.; Iatagan, M.; Geamanu, M.; Ștefanescu, R.; Dijmarescu, I. Deep Learning-Assisted Smart Process Planning, Robotic Wireless Sensor Networks, and Geospatial Big Data Management Algorithms in the Internet of Manufacturing Things. *ISPRS Int. J. Geo Inf.* **2022**, *11*, 277. [CrossRef]
93. Pelau, C.; Dabija, D.-C.; Ene, I. What Makes an AI Device Human-Like? The Role of Interaction Quality, Empathy and Perceived Psychological Anthropomorphic Characteristics in the Acceptance of Artificial Intelligence in the Service Industry. *Comput. Hum. Behav.* **2021**, *122*, 106855. [CrossRef]
94. Frajtova Michalikova, K.; Suler, P.; Robinson, R. Virtual Hiring and Training Processes in the Metaverse: Remote Work Apps, Sensory Algorithmic Devices, and Decision Intelligence and Modeling. *Psychosociological Issues Hum. Resour. Manag.* **2022**, *10*, 50–63. [CrossRef]
95. Cazazian, R. Blockchain Technology Adoption in Artificial Intelligence-based Digital Financial Services, Accounting Information Systems, and Audit Quality Control. *Rev. Contemp. Philos.* **2022**, *21*, 55–71. [CrossRef]
96. Nica, E.; Kliestik, T.; Valaskova, K.; Sabie, O.-M. The Economics of the Metaverse: Immersive Virtual Technologies, Consumer Digital Engagement, and Augmented Reality Shopping Experience. *Smart Gov.* **2022**, *1*, 21–34. [CrossRef]
97. Valaskova, K.; Machova, V.; Lewis, E. Virtual Marketplace Dynamics Data, Spatial Analytics, and Customer Engagement Tools in a Real-Time Interoperable Decentralized Metaverse. *Linguist. Philos. Investig.* **2022**, *21*, 105–120. [CrossRef]
98. Pocol, C.B.; Stanca, L.; Dabija, D.-C.; Pop, I.D.; Mișcoiu, S. Knowledge Co-creation and Sustainable Education in the Labor Market-Driven University–Business Environment. *Front. Environ. Sci.* **2022**, *10*, 781075. [CrossRef]
99. Kral, P.; Janoskova, K.; Dawson, A. Virtual Skill Acquisition, Remote Working Tools, and Employee Engagement and Retention on Blockchain-based Metaverse Platforms. *Psychosociological Issues Hum. Resour. Manag.* **2022**, *10*, 92–105. [CrossRef]
100. Andronie, M.; Lazaroiu, G.; Ștefanescu, R.; Uța, C.; Dijmarescu, I. Sustainable, Smart, and Sensing Technologies for Cyber-Physical Manufacturing Systems: A Systematic Literature Review. *Sustainability* **2021**, *13*, 5495. [CrossRef]
101. Peters, M.A. Poststructuralism and the Post-Marxist Critique of Knowledge Capitalism: A Personal Account. *Rev. Contemp. Philos.* **2022**, *21*, 21–37. [CrossRef]
102. Vatamanescu, E.-M.; Alexandru, V.-A.; Mitan, A.; Dabija, D.-C. From the Deliberate Managerial Strategy towards International Business Performance: A Psychic Distance vs. Global Mindset Approach. *Syst. Res. Behav. Sci.* **2020**, *37*, 374–387. [CrossRef]
103. Durana, P.; Krulicky, T.; Taylor, E. Working in the Metaverse: Virtual Recruitment, Cognitive Analytics Management, and Immersive Visualization Systems. *Psychosociological Issues Hum. Resour. Manag.* **2022**, *10*, 135–148. [CrossRef]
104. Poliak, M.; Jurecki, R.; Buckner, K. Autonomous Vehicle Routing and Navigation, Mobility Simulation and Traffic Flow Prediction Tools, and Deep Learning Object Detection Technology in Smart Sustainable Urban Transport Systems. *Contemp. Read. Law Soc. Justice* **2022**, *14*, 25–40. [CrossRef]
105. Kovacova, M.; Novak, A.; Machova, V.; Carey, B. 3D Virtual Simulation Technology, Digital Twin Modeling, and Geospatial Data Mining in Smart Sustainable City Governance and Management. *Geopolit. Hist. Int. Relat.* **2022**, *14*, 43–58. [CrossRef]
106. Blake, R.; Frajtova Michalikova, K. Deep Learning-based Sensing Technologies, Artificial Intelligence-based Decision-Making Algorithms, and Big Geospatial Data Analytics in Cognitive Internet of Things. *Anal. Metaphys.* **2021**, *20*, 159–173. [CrossRef]
107. Kliestik, T.; Musa, H.; Machova, V.; Rice, L. Remote Sensing Data Fusion Techniques, Autonomous Vehicle Driving Perception Algorithms, and Mobility Simulation Tools in Smart Transportation Systems. *Contemp. Read. Law Soc. Justice* **2022**, *14*, 137–152. [CrossRef]
108. Zvarikova, K.; Horak, J.; Downs, S. Digital Twin Algorithms, Smart City Technologies, and 3D Spatio-Temporal Simulations in Virtual Urban Environments. *Geopolit. Hist. Int. Relat.* **2022**, *14*, 139–154. [CrossRef]
109. Lăzăroiu, G.; Ionescu, L.; Andronie, M.; Dijmărescu, I. Sustainability Management and Performance in the Urban Corporate Economy: A Systematic Literature Review. *Sustainability* **2020**, *12*, 7705. [CrossRef]
110. Carey, B. Geospatial Data Mining, Computer Vision-based Lane Detection and Object Localization Algorithms, and Mapping and Navigation Tools in Smart Urban Mobility Systems. *Contemp. Read. Law Soc. Justice* **2022**, *14*, 121–136. [CrossRef]
111. Kliestik, T.; Poliak, M.; Popescu, G.H. Digital Twin Simulation and Modeling Tools, Computer Vision Algorithms, and Urban Sensing Technologies in Immersive 3D Environments. *Geopolit. Hist. Int. Relat.* **2022**, *14*, 9–25. [CrossRef]
112. Andronie, M.; Lăzăroiu, G.; Iatagan, M.; Hurloiu, I.; Dijmărescu, I. Sustainable Cyber-Physical Production Systems in Big Data-Driven Smart Urban Economy: A Systematic Literature Review. *Sustainability* **2021**, *13*, 751. [CrossRef]
113. Konecny, V.; Jaśkiewicz, M.; Downs, S. Motion Planning and Object Recognition Algorithms, Vehicle Navigation and Collision Avoidance Technologies, and Geospatial Data Visualization in Network Connectivity Systems. *Contemp. Read. Law Soc. Justice* **2022**, *14*, 89–104. [CrossRef]
114. Durana, P.; Krastev, V.; Buckner, K. Digital Twin Modeling, Multi-Sensor Fusion Technology, and Data Mining Algorithms in Cloud and Edge Computing-based Smart City Environments. *Geopolit. Hist. Int. Relat.* **2022**, *14*, 91–106. [CrossRef]
115. Lăzăroiu, G.; Pera, A.; Ștefănescu-Mihăilă, R.O.; Mircică, N.; Neguriță, O. Can Neuroscience Assist Us in Constructing Better Patterns of Economic Decision-Making? *Front. Behav. Neurosci.* **2017**, *11*, 188. [CrossRef]

mathematics

MDPI

Article

RecessionRisk^{+}: A Novel Recession Risk Model with Applications to the Solvency II Framework and Recession Crises Forecasting

Jacopo Giacomelli [1,2,*,†] and Luca Passalacqua [2]

1 SACE S.p.A., Piazza Poli 42, 00187 Rome, Italy
2 Department of Statistics, Sapienza University of Rome, Viale Regina Elena 295, 00161 Rome, Italy; luca.passalacqua@uniroma1.it
* Correspondence: j.giacomelli@sace.it
† The views and opinions expressed in this article are those of the author and do not necessarily reflect the official policy or position of SACE S.p.A.

Abstract: The Solvency II regulatory framework requires European insurance companies to guarantee their solvability and stability by retaining enough Own Funds to cover future unexpected losses at a given confidence level. A Standard Formula approach is provided to estimate the capital requirement needed. Still, Solvency II allows internal methodologies to quantify the capital absorption arising from specific risk types or even to replace the Standard Formula with a full internal model. This work proposes a new internal model approach to measure the Catastrophe Recession Risk. The Recession Risk implies a mandatory capital absorption component for the insurance companies operating in the credit and suretyship business. The proposed model is based on the CreditRisk^{+} model and designed to behave countercyclically, aligning with the original intent of the European supervisory authority when first introducing this risk into the Solvency II risks' taxonomy. Additionally, the model is applied to define an index for monitoring future recession crises based on the time series of past default rates.

Keywords: credit risk; recession; risk management; Solvency II; suretyship

MSC: 91G40; 91G70; 62P05

1. Introduction

The business cycle—also commonly referred to as the economic cycle—is the concept used to describe the fluctuating behavior of an economy, featuring alternating expansion and contraction phases [1,2]. The word "recession" refers to the contraction phases in slightly different meanings depending on the context, as a rigorous, universally accepted definition of recession is missing to date [3].

A popular definition of a recession is two consecutive quarters of decline in a country's real (i.e., inflation-adjusted) gross domestic product [3–5]. Recession may also refer to the whole period between a peak of economic activity and its subsequent trough or, even more generically, to a period of decline in economic activity [6]. However, every accepted definition requires that the considered period of economic contraction satisfies, to some extent, the criteria of depth, diffusion, and duration of the observed decline to be legitimately called a recession.

Recessions are characterized by a general increase in the likelihood of defaults (see, e.g., refs. [7,8] and the definition of Moody's scenario "S3" in [9]). Thus, the frequency of default and insolvency events observed during a recession is significantly higher than average. In this work, "Recession Risk" refers to the potential unexpected losses experienced by an insurance company, a bank, or other financial institution in such an adverse scenario, due to the increased number of insolvent counterparties.

In the Solvency II regulatory framework [10–12], Recession Risk belongs to the broader set of Catastrophe Non-Life risks. These components of the Solvency Capital Requirement (SCR) are generally introduced to increase the capital requirement contribution from the Premium Risk, which measures the unexpected losses arising from future claims that an insurance company must indemnify during the following year.

In the Solvency II taxonomy of risks, most catastrophe risks are associated with claim generator events that are very specific in nature and, thus, may not be adequately considered in the corresponding Premium Risk measure. For example, concerning the damages to property insurance, the economic losses generated by an earthquake are classified as a catastrophic event, as opposed to a generic repair needed by an insured building, whose probability and severity are accounted for in the Premium Risk measure.

On the other hand, there is no clear distinction between the claims generated by a recession event and the claims considered in the corresponding Premium Risk measure. The Recession Risk, together with the Catastrophe Default Risk, is specifically considered in the SCR measure of the Solvency 2 Line of Business 9 (credit and suretyship insurance—C&S) [11]. The C&S Premium Risk, Catastrophe Default Risk, and Recession Risk are introduced to measure the unexpected losses arising from insured events of default, insolvency, and breach of the undertaking [13–15]. The Solvency II Calibration Paper clarifies that the original role of the Recession Risk measure was to dampen the procyclical nature of the LoB C&S. The European regulator EIOPA explicitly declared the original purpose of this submodule in the "Solvency II Calibration Paper" (see [16] §3.1153 p. 312), where a counter-cyclical dampening mechanism has been introduced—to be abandoned later in the 2015 release of Standard Formula technical specifications [11]: "[. . .] *This mechanism aims to ensure that at the peak of the cycle (low failure rates), the* $SCR_{CAT\_recession\_ratio\_net}$ *shall reach its highest value and C&S undertakings shall be required to have enough own funds to cover a higher SCR. On the other hand, at the trough of the cycle, SCR will be at its lowest value so that own funds will be released. In other words, as undertakings face a harder net claims ratio due to an increase of failure rates, the SCR decreases*".

Since the event type considered in C&S Recession Risk is the same as the Premium Risk, double-counting is possible if not prevented in calibration. Indeed, this is the view expressed by ICISA in 2018 [17], answering to EIOPA's "*Second, set of advice to the European Commission on specific items in the Solvency II Delegated Regulation*" [18]. In its answer, ICISA pointed out that the C&S Premium Risk parameter's calibration has been made without excluding recession events from the considered data. Thus, default events arising during a recession event should already be included in the Premium Risk measure, which aims to quantify the following year's unexpected losses at the 99.5% confidence level.

EIOPA did not consider the objection of ICISA in the review of the Standard Formula [12]. Further, the Recession Risk SCR measure was simplified, losing the dampening feature that originally motivated its introduction. Nonetheless, Solvency II allows both the calibration of the Standard Formula parameters based on an insurance company's dataset (the so-called Undertaking Specific Parameters) and the replacement of specific modules and submodules of the Standard Formula with a Partial Internal Model (PIM), given that both these possibilities are subject to the preliminary verification and formal approval by a National Supervisory Authority [10].

In this work, we propose a model that describes and measures the Recession Risk based on the original EIOPA's perspective on this risk. The model can be considered a PIM approach to measure the C&S Recession Risk SCR, providing a coherent C&S Premium Risk SCR as well. These two SCR components restore the counter-cyclical dampening mechanism initially considered by EIOPA but with the granularity and realism required for a PIM. Additionally, the model can be calibrated entirely based on the historical dataset typically available to a C&S insurance company. Datasets adequate for model calibration should also be available for banks, financial institutions, and supervisory authorities. Indeed, the application of the proposed model is not limited to the insurance sector; it can also be used to define a new set of indexes and anticipate future recession events. Both the applications—SCR computation and crisis monitoring—are discussed throughout the remainder of this work.

As the name proposed for the model suggests, our approach is inspired by the classic CreditRisk^{+} model [19–22], which has been shown to describe well the underwriting risk both in credit insurance [23] and suretyship [13]. However, we have based this work on the modified version proposed in [22] and extended the modifications, resulting in a novel approach specifically designed to address the Recession Risk measurement. Overall, the one-year unexpected loss level corresponding to the C&S Premium Risk SCR is measured by a point-in-time (PIT) [24,25] loss distribution calibrated according to the current phase of the economic cycle. At the same time, the Recession Risk SCR is a function of the difference between Value at Risk (VaR) measures, defined at the same quantile of the PIT distribution mentioned above and a through-the-cycle (TTC) loss distribution, with the latter being calibrated including previously observed recession phases. In this way, at the peak of the economic cycle, we have the maximum difference between the two measures. In contrast, at the trough of the cycle, the PIT VaR results remarkably close to or even greater than the TTC VaR, decreasing Recession Risk and releasing own funds. The two distributions involved in the Premium and Recession risks measurement are naturally implied by the same set of assumptions, featuring a short-term autocorrelation in the probability of default time series to distinguish the one-year PIT loss distribution from the corresponding TTC distribution.

CreditRisk^{+} is a uniperiodal model where the realization of the latent market factors per scenario affects the local deviation of each subject's probability of default from its TTC level. Although probabilities can be instantaneously shocked, they cannot be described through processes, and the model does not allow for the inference of trends from historical data. Conversely, RecessionRisk^{+} features a consistent multiperiod framework that describes market factors as processes over time. Additionally, the autocorrelation mentioned earlier introduces a memory effect within each process, leading to the possibility of modeling not only the current deviation from the TTC probability but also the existence of a trend towards or away from the through of the economic cycle. These two enhancements of the original CreditRisk^{+} framework allow the model to represent the approach to recession periods and, thus, to measure the recession risk.

Many studies are available concerning the modeling of recessions as a phase of the economic cycle (see, e.g., ref. [26]). In this perspective, the concept of the financial cycle and its relation with the business cycle has also been deeply investigated [27,28]. According to this macro-economic approach, the possible occurrence of a recession is seen as a phase of the interrelated dynamics of financial markets and the real economy. However, in this framework, default probabilities are not explanatory variables and recessions are mainly measured and modeled by considering GDP level, employment rate, interest rates, and other macroeconomic and financial variables as risk drivers (see, e.g., refs. [29,30]). One of the first models of this type is the "classic" CreditPortfolioView [31,32], where the probabilities of default are assumed to be a function of macro-economic variables, represented through autoregressive processes. Although not explicitly focused on recessions, these models allow to map the business cycle dynamics investigated in the aforementioned macro-economic literature to the default probabilities, thus being able to project an increasing number of default events in the presence of adverse macro-economic scenarios. The drawback of such a composite approach to managing Recession Risk is the vast data requirements to calibrate and validate the models, whose assumptions could need a high level of maintenance depending on the context. Furthermore, contagion models (see, e.g., ref. [33]) improve the macro-economic-inspired models by adding a further layer of complexity, based on the micro-economic description of direct connections between firms.

The proposed model is not comparable to any of these two classes of models, as it can be considered a reduced form model, given that the multivariate hazard rates dynamics is modeled by itself, without any explicit assumption concerning an underlying dependency on macro-economic or financial quantities. The elementary idea behind this method is that, given the cyclic nature of a considered economy, the comparison of the PIT and TTC hazard rate distributions (in particular of their tails) provides information concerning the likelihood of a recession event and, thus, a measure of Recession Risk. Although the banking research in the literature vastly discusses the PIT and TTC calibration approaches (see, e.g., ref. [34]), their specific and joint application to a multivariate Recession Risk model is new and worth investigating.

The work is organized as follows. Section 2 defines the proposed RecessionRisk$^+$ model, introducing first the CreditRisk$^+$ model to which our approach is inspired and its extension mentioned above. The same section provides a complete calibration method based on historical time series and describes the two applications of the model proposed in this work—the construction of a multivariate index to monitor the likelihood of future recession periods and the Recession Risk measurement in the Solvency II regulatory framework. The following Section 3 applies the model to a real dataset concerning the non-performing loans time series observed in Italy's Banking System. Both the applications proposed in the previous Section 2 are implemented based on the considered dataset. Finally, Section 4 summarizes the main results of this work, also discussing the possible developments of this topic that are worth being investigated further.

2. Models and Methods

As anticipated in Section 1, in this work, we consider the extended CreditRisk$^+$ framework developed in [22] and develop it further to measure the additional credit risk that may arise in the case of a recession crisis. To do so, in Section 2.1, the main assumptions underlying the original CreditRisk$^+$ model are briefly summarized. In the following Section 2.2, we describe the extended framework that was previously developed. The further model extension introduced in this work is reported in Section 2.3, and a full calibration method for the new model is provided in Section 2.4. The proposed model is then utilized to define a Recession Risk forecasting index in Section 2.5. Finally, in Section 2.6, we apply the model to introduce a measure of the Recession Risk that copes with the Solvency II semantics.

2.1. The Original CreditRisk$^+$ Model

The CreditRisk$^+$ model ("$CR+$" hereafter) was initially developed by Tom Wilde in 1997 and first disclosed by CSFB in 1998. $CR+$ is a well-known portfolio model, widely used across the financial industry in the recent three decades to quantify the distribution of future losses in t generated by a portfolio of risky debtors up to a projection time $T > t$.

The model follows a frequency-severity paradigm, typical in actuarial non-life models. Indeed, the likelihood of a future default event is described separately from the severity (i.e., the so-called "exposure at default", also EAD, in a bank's perspective) of the same event. In its original formulation, $CR+$ assumes that EAD_i is fixed and deterministic per each i-th debtor.

Default events are assumed to be conditionally independent, although a dependency structure is imposed among the debtors' default probabilities (abbreviated as PDs hereafter). This structure is represented by introducing a set of independent, Gamma-distributed random variables, also known as "market factors". Each PD depends linearly on the market factors through a set of weights, commonly addressed as factor loadings, which, in principle, may be different per debtor in the considered portfolio. The market factors' realization represents the future market condition that a debtor experiences across the time interval $(t, T]$: a wide shock in the Gamma latent factors implies that the debtors' creditworthiness is generally worsened and causes the projected default events, although mutually independent, to increase in number.

Potential issues in PDs' definition, assuming Gamma-distributed shocks to be applied to variables defined in the $[0, 1]$ interval, are prevented by approximating the Bernoulli variable that describes each debtor's default event with a Poisson variable. This substitution extends the domain of the distribution's parameter from $[0, 1]$ to $\mathbb{R}_+$. Further, it is numerically sound if the parameter's expected value is "small" compared with 1. This is easily verified, recalling that the probability generating function of a Bernoulli r.v. can be read as the first-order approximation of the Poisson r.v.'s p.g.f. with the same parameter. Such a substitution, together with the assumption above concerning a deterministic EAD, enables a semi-analytical solution of the model. The original $CR+$ set of assumptions concerning the PDs can be summarized as follows: [19–21].

Assumption 1 ($CR+$ distributional assumption)**.** *Given a time horizon $(t, T]$ and a set of N risky debtors, the number Y_i of insolvency events generated by each i-th debtor over $(t, T]$ is distributed as follows:*

$$Y_i \sim \text{Poisson}(p_i(\mathbf{\Gamma})), \quad p_i(\mathbf{\Gamma}) := q_i \cdot \left(\omega_{i0} + \sum_{k=1}^{K} \omega_{ik}\Gamma_k\right) \tag{1}$$

where $\mathbf{\Gamma} = (\Gamma_1 \dots \Gamma_K) \in \mathbb{R}_+^K$ is an array of independent r.v.'s such that

$$\Gamma_k \sim \text{Gamma}\left(\beta_k^{-1}, \beta_k\right), \quad \beta_k \in \mathbb{R}_+ \tag{2}$$

and the factor loadings ω_{ik} *are supposed to be all non-negative and to sum up to unity:*

$$\begin{aligned} &\omega_{ik} \geq 0, \qquad i = 1, \dots, N, \quad k = 0, \dots, K, \\ &\sum_{k=0}^{K} \omega_{ik} = 1, \quad i = 1, \dots, N. \end{aligned} \tag{3}$$

2.2. The Multiperiodal Autocorrelated CreditRisk$^+$ Extension

As discussed above in Section 2.1, the original $CR+$ model is defined in a uniperiodal framework. The model aims to simulate the terminal loss distribution in t that is observed in $T > t$. Each scenario terminal loss is the sum of the losses generated by the considered set of risky debtors onto the future time horizon $(t, T]$. The classic model's calibration procedure infers the parameters' values from the historical time series defined with a sampling period equal to the duration of the simulation time horizon (e.g., if $T - t = 1$ year, the model's calibration considers yearly default rates time series as well). Namely, the model is defined considering just a single time scale.

Indeed, as shown in [22], it is possible and useful to generalize the $CR+$ model from the original single-time-scale framework to become a multiple-time-scale model. Such an extension is especially beneficial in terms of error reduction when calibrating the $CR+$ parameters based on, e.g., monthly or quarterly default rates time series, to apply the calibrated model for simulating a terminal loss distribution defined over a one-year time horizon. Indeed, calibrating the model based on data sampled with a given time scale (e.g., one month) shorter than the one considered for the model application (e.g., one year) leads to a more efficient inference process, provided that Assumption 1 is satisfied consistently at both the considered time scales, returning estimates of the model's parameters much more precisely than the ones attainable with a single-time-scale calibration strategy. This is the case, as proved in [22], under different configurations.

In the following, we consider only a specific configuration of the generalized $CR+$ model among the ones discussed in [22]. The remainder of this subsection outlines the chosen configuration, upon which we define the model proposed in this work.

First, we recover the Bernoulli representation of each debtor by introducing the r.v.

$$\widetilde{Y}_i := \mathbb{1}_{Y_i > 0}. \tag{4}$$

Both the r.v. Y_i and its distribution parameter $p_i(\mathbf{\Gamma})$ can take values larger than 1. This is formally correct, given that $Y_i \sim \text{Poisson}(p_i(\mathbf{\Gamma}))$, despite not coping with the representation of absorbing events, which can occur at most once by definition. The so-called "Poisson approximation", introduced by substituting $\mathbb{I}_i$ with Y_i, is numerically sound as q_i approaches to zero—a condition that is well fulfilled in most real world relevant cases. Indeed, Assumption 1 implies that $\widetilde{Y}_i|\mathbf{\Gamma} \sim \text{Bernoulli}(\tilde{p}_i(\mathbf{\Gamma}))$, where the distribution parameter is

$$\tilde{p}_i(\mathbf{\Gamma}) = \text{Prob}(Y_i > 0|\mathbf{\Gamma}) = 1 - \exp\left[-q_i\left(\omega_{i0} + \sum_{k=1}^{K} \omega_{ik}\,\Gamma_k\right)\right]. \tag{5}$$

Representing the i-th debtor through $\widetilde{Y}_i$ instead of Y_i changes Assumption 1 into the following:

Assumption 2 (Modified $CR+$ distributional assumption)**.** *Given a time horizon $(t, T]$ and a set of N risky debtors, the number of insolvency events generated by each i-th debtor over $(t, T]$ is represented by the r.v. $\widetilde{Y}_i \sim$ Bernoulli$(\tilde{p}_i(\mathbf{\Gamma}))$, where the distribution parameter $\tilde{p}_i(\mathbf{\Gamma})$ satisfies* (5)*. Assumptions on market factors' array $\mathbf{\Gamma}$ and factor loadings' matrix $\mathbf{\Omega}$ remain the same as stated in Assumption 1.*

Assumption 2 is well posed also in case it is required to hold at different time scales, resulting into a self-consistent multiperiodal model. Before describing the multiple-time-scale version of $CR+$, let us introduce two simplifications that will be maintained in the remainder of this work.

First, each debtor belongs to a cluster $c_h : h \in \{1 \dots H\}$. All the debtors belonging to the same cluster c_h are described by the same set of parameters, allowing the re-indexing of the parameters themselves. Namely, it holds

$$\begin{matrix} q_i = q_{i'} \equiv q_h; \\ \omega_{ik} = \omega_{i'k} \equiv \omega_{hk} \end{matrix} \qquad \forall i, i' \in c_h. \tag{6}$$

This choice is relevant to practical purposes. Indeed, it reduces the number of parameters significantly, making the model calibration more achievable.

Further, the partition $\{t \equiv t_0, t_1, \dots, t_m \equiv T\}$ of the time interval $(t, T]$ needed in the following is assumed to always be uniform:

$$\frac{t_j - t_{j-1}}{T - t} = \frac{1}{m} \quad \forall j \in \{1 \dots m\} \tag{7}$$

Although not strictly needed, these two simplifications cope with real-world applications of the model and make the results reported in this work easier to be proved and represented. Equipped with Equations (6) and (7), we are now ready to introduce the multiperiodal framework mentioned above.

Assumption 3 (Modified $CR+$ parameters at different time scales)**.** *Let $t \equiv t_0, t_1, \dots, t_m \equiv T$ be a partition of the time interval $(t, T]$. Let Assumption 2 be verified by each j-th interval $(t_{j-1}, t_j]$, where $\widetilde{Y}_i^{(j)}$ is the r.v. representing the i-th risk observed during the j-th interval. Further, the associated set $\{q_i^{(j)}; \gamma^{(j)}; \Omega^{(j)}\}$ of parameters and market factors satisfies the following for each $j \in \{1 \dots m\}$:*

$$q_i^{(j)} = \frac{q_i}{m}, \tag{8}$$

$$\gamma_k^{(j)} \sim \text{Gamma}\left(\sigma_k^{-2}\,\xi_k^{-2}\,m^{-1}, \sigma_k^2\,\xi_k^2\,m\right), \tag{9}$$

$$\Omega^{(j)} = \Omega, \tag{10}$$

where Ω *is the factor loadings' matrix and*

$$\xi_k := \left[1 + 2\sum_{x=1}^{m} \rho_{xk}\left(1 - \frac{x}{m}\right)\right]^{-\frac{1}{2}}. \tag{11}$$

Furthermore, $\rho_{xk} \in [0,1]$ *is assumed to be a time-invariant ACF:*

$$\boldsymbol{cov}\left(\gamma_k^{(j)}, \gamma_k^{(j+x)}\right) = \varrho_{xk}\boldsymbol{var}\left(\gamma_k^{(j)}\right), \tag{12}$$

such that the following closure property with respect to the addition holds:

$$\sum_{j=1}^{m} \gamma_k^{(j)} \sim \text{Gamma}(\alpha_k, \beta_k) \tag{13}$$

for a proper shape and scale parameters' specification α_k, β_k*.*

Equation (13) is satisfied by construction in case of summating independent Gamma random variables with the same scale parameters. The same is shown to hold with a good degree of approximation in case the Gamma variables are weakly correlated [35]. In [22], it is proved that the framework introduced in Assumption 3, (i.e., multiperiodal, short-time-scaled, autocorrelated) is consistent with the modified $CR+$ model defined in Assumption 2 (i.e., uniperiodal, long-time-scaled). For readers' convenience, this result is summarized below with a notation that is lighter than the original one.

Proposition 1 (Consistency of the modified $CR+$ model)**.** *Assumption 3 implies Assumption 2.*

Proof. According to Assumption 3, Equation (5) holds for each interval $(t_{j-1}, t_j]$ ($j = 1 \ldots m$). Thus, Equations (8) and (10) imply that the survival function $\mathcal{S}_i(t,T)$ of the i-th debtor over the interval $(t,T]$ can be written as follows:

$$\begin{aligned}
-\ln \mathcal{S}_i(t,T) &= -\sum_{j=1}^{m} \ln \mathcal{S}_i(t_{j-1}, t_j) \\
&= \sum_{j=1}^{m} \frac{q_i}{m}\left(\omega_{i0} + \sum_{k=1}^{K} \omega_{ik}\,\gamma_k^{(j)}\right) \\
&= q_i\left(\omega_{i0} + \sum_{k=1}^{K} \omega_{ik}\,\frac{1}{m}\sum_{j=1}^{m} \gamma_k^{(j)}\right)
\end{aligned}$$

The proof is completed by showing that $\frac{1}{m}\sum_{j=1}^{m} \gamma_k^{(j)} =: \Sigma\gamma_k \equiv \Gamma_k$. Since the Gamma distribution is fully specified by its first two central moments, Equation (13) implies that the equivalence holds if $\mathbf{E}[\Sigma\gamma_k] = \mathbf{E}[\Gamma_k]$ and $\mathbf{var}[\Sigma\gamma_k] = \mathbf{var}[\Gamma_k]$. The expectations' requirement is immediately verified, given that $\mathbf{E}\left[\gamma_k^{(j)}\right] = \mathbf{E}[\Gamma_k] = 1$ for $j = 1 \ldots m$. As for the variances, Equations (11) and (12) imply

$$\begin{aligned}
\mathbf{var}\left[\sum_{j=1}^{m} \gamma_k^{(j)}\right] &= \sum_{j=1}^{m} \mathbf{var}\left[\gamma_k^{(j)}\right] + \sum_{j=1}^{m}\sum_{j' \neq j} \mathbf{cov}\left[\gamma_k^{(j)}, \gamma_k^{(j')}\right] \\
&= \mathbf{var}\left[\gamma_k^{(1)}\right] \underbrace{\left(m + 2\sum_{x=1}^{m-1}(m-x)\varrho_{xk}\right)}_{=m\xi_k^{-2}} = m^2\sigma_k^2.
\end{aligned}$$

The equivalence of the first two central moments and the closure assumption implies that the shape and scale parameters that solve Equation (13) in Assumption 3 are the same regarding Γ_k in Assumption 2, completing the proof. $\square$

It is worth remarking that Equation (13) is assumed to be true but, actually, it is not satisfied by each ϱ_{xk} form that could be chosen in Equation (12). As stated above, the trivial case $\varrho_{xk} = 0$ verifies Equation (13). Further, given that the ACF obeys a power law, the sum-variable $\Sigma\Gamma_k$ is known to be approximately Gamma-distributed [35], while other ϱ_{xk} specifications imply the sum to be distributed differently [36,37]. Thus, for practical purposes, we complete the modified $CR+$ framework by introducing the following assumption:

Assumption 4 (Modified $CR+$ ACF)**.** *The time-invariant ACF ϱ_{xk} considered in Equation* (12) *obeys the power law*

$$\varrho_{xk} = \rho_k^{|x|},$$

where $\rho_k \in (0,1)$.

2.3. The Double-Distribution Framework

The two-time-scale framework introduced so far can be utilized to calibrate the modified $CR+$ model from the "short" time scale $\delta := (T-t)/m$, and to project loss scenarios onto the "long" time scale $\Delta := T - t$, as demonstrated in [22,23]. In this subsection, the framework is completed to use both time scales for projection purposes.

To that end, we provide a notation representing the hazard rate processes defined onto the two time scales Δ and δ.

$$\begin{aligned} \lambda_h(t) &:= -\ln \mathcal{S}_h(t, t+\delta), \\ \Lambda_h(t) &:= -\ln \mathcal{S}_h(t, t+\Delta). \end{aligned}$$

We are now using the index "h" instead of "i", given that we are assuming that the i-th debtor is fully specified by the h-th cluster to which it belongs. Assumptions 2 and 3 and Proposition 1 imply the following.

$$\lambda_h(t) = \frac{q_h}{m}\left[\omega_{h0} + \sum_{k=1}^{K} \omega_{hk}\,\gamma_k(t)\right], \tag{14}$$

$$\begin{aligned} \Lambda_h(t) &= q_h\left[\omega_{h0} + \sum_{k=1}^{K} \omega_{hk}\,\Gamma_k(t)\right] \\ &= q_h\left[\omega_{h0} + \sum_{k=1}^{K} \omega_{hk}\,\frac{1}{m}\sum_{x=0}^{m-1} \gamma_k(t+x\delta)\right] \\ &= \sum_{x=0}^{m-1} \lambda_h(t+x\delta). \end{aligned} \tag{15}$$

The processes $\lambda_h(t)$ and $\Lambda_h(t)$ displayed above are unconditioned, in the sense that they do not take into account the past value of the latent market factors' array $\gamma(t-\delta) = \overline{\gamma} \in \mathbb{R}^K$. From an economic perspective, they can be regarded as *through-the-cycle* distributions of the h-th hazard rate onto the two time scales Δ and δ, because they are evaluated without considering which phase of the macroeconomic cycle is occurring at t.

On the other hand, the corresponding conditioned processes, $\lambda_i^x(t,\overline{\gamma})$ and $\Lambda_i^\star(t,\overline{\gamma})$, may be defined as well.

$$\begin{aligned} \lambda_h^x(t,\overline{\gamma}) &:= -\ln \mathcal{S}_h(t,t+\delta)\,|\,\gamma(t-x\delta) = \overline{\gamma}, \quad x \in \mathbb{N}, \\ \Lambda_h^\star(t,\overline{\gamma}) &:= -\ln \mathcal{S}_h(t,t+\Delta)\,|\,\gamma(t-\delta) = \overline{\gamma}. \end{aligned}$$

These conditioned distributions can be thought of as *point-in-time* from an economic perspective, as they depend on the most recently occurred market state $\overline{\gamma}$. Thus, they are influenced by the current phase of the macroeconomic cycle.

We define $\Lambda^{\star}_{h}$ evaluated at time t to be conditioned only on the value of the δ-time-scale market factors at $t-\delta$. At the same time, the definition of λ^{x}_{h} allows us to utilize arbitrarily past values $\gamma(t-x\delta)$ of the market factors.

The set of assumptions stated so far enables us to write closed-form expressions for $\lambda^{x}_{h}(t,\overline{\gamma})$ and $\Lambda^{\star}_{h}(t,\overline{\gamma})$, provided that $\sigma_k \ll 1$ ($k = 1\dots K$), as shown by the following proposition.

Proposition 2 (point-in-time hazard rate distributions)**.** *Assumptions 3 and 4 imply that*

$$\lim_{\sigma\to\underline{0}^+} \lambda^{x}_{h}(t,\overline{\gamma}) = \frac{q_h}{m}\left[\omega_{h0} + \sum_{k=1}^{K} \omega_{hk}\left[(\gamma_k(t)-1)\sqrt{1-\rho_k^{2|x|}} + 1 + \rho_k^{|x|}(\overline{\gamma}_k - 1)\right]\right] \tag{16}$$

$$\lim_{\sigma\to\underline{0}^+} \Lambda^{\star}_{h}(t,\overline{\gamma}) = \sum_{x=0}^{m-1} \lambda^{1}_{i}(t_x,\gamma(t_{x-1})); \quad \gamma(t_{-1}) = \overline{\gamma} \tag{17}$$

where $\sigma := (\sigma_1,\dots,\sigma_K)$ *and* $t_x := t + x\delta$.

Proof. In the limit $\sigma \to \underline{0}^+$, we can approximate the Gamma-distributed market factors $\gamma \in \mathbb{R}^K_+$ as being normally distributed, as shown below in Proposition A1. Doing so, we can estimate the conditioned expected value and variance of $\gamma(t)|\gamma(t-x\delta)$ ($x \in \mathbb{N}$) as follows:

$$\mathbf{E}[\gamma_k(t)|\gamma_k(t-x\delta)] = 1 + \rho_k^{|x|}[\gamma_k(t-x\delta) - 1], \tag{18}$$

$$\mathbf{var}[\gamma_k(t)|\gamma_k(t-x\delta)] = m\xi_k^2\sigma_k^2\left(1-\rho_k^{2|x|}\right). \tag{19}$$

Equations (18) and (19) are immediately implied by applying Proposition A2 to the modified $CR+$ model outlined in Assumptions 3 and 4.

With specific reference to the short time scale δ, we now need to distinguish between the unconditioned market factor $\gamma_k(t)$, which is distributed as stated in Equation (9), and the conditioned market factor $\gamma^{x}_{k}(t,\overline{\gamma}_k)$, whose distribution follows from Equations (18) and (19):

$$\begin{aligned}
\gamma^{x}_{k}(t,\overline{\gamma}_k) &:= \gamma_k(t)|_{\gamma_k(t-x\delta)=\overline{\gamma}_k} \\
&= (\gamma_k(t) - \mathbf{E}[\gamma_k(t)])\sqrt{\frac{\mathbf{var}[\gamma_k(t)|\gamma_k(t-x\delta)=\overline{\gamma}_k]}{\mathbf{var}[\gamma_k(t)]}} \\
&+ \mathbf{E}[\gamma_k(t)|\gamma_k(t-x\delta)=\overline{\gamma}_k] \\
&= (\gamma_k(t)-1)\sqrt{1-\rho_k^{2|x|}} + 1 + \rho_k^{|x|}(\overline{\gamma}_k - 1)
\end{aligned}$$

Since Assumption 3 allows us to write the conditioned hazard rate onto the short time scale as

$$\lambda^{x}_{h}(t,\overline{\gamma}) = \frac{q_h}{m}\left[\omega_{h0} + \sum_{k=1}^{K}\omega_{hk}\,\gamma^{x}_{k}(t,\overline{\gamma}_k)\right],$$

the functional form of $\gamma^{x}_{k}(t,\overline{\gamma}_k)$ determined above implies Equation (16).

Proposition A1 and Assumption 4 are consistent with each $\gamma_k(t)$ ($k = 1\dots K; t \in \delta\mathbb{N}$) being an $AR(1)$ process in the limit $\sigma_k \to \underline{0}^+$. Thus, only conditioning each $\lambda_h(t_x)$ to $\gamma(t_{x-1})$ is needed to cope with the assumed ACF, implying (17) and completing the proof. □

It is worth noticing that—generally—we need to infer the value $\overline{\gamma}$ taken by the latent market factors' array $\gamma(t-\delta)$.

Remark 1. *We can avoid estimating* $\overline{\gamma}$ *only in the special case* $\rho_k = \rho$ ($k = 1\dots K$)*, which has a modest interest from a practical perspective. Indeed, in case of uniform* ρ *across the market factors, Equation* (16) *can be written as a linear combination of* $\lambda_h(t)$ *and* $\lambda_h(t-\delta)$.

The RecessionRisk$^+$ model consists of the two multivariate hazard rate distributions introduced to represent and compare the PIT and the TTC one-year tail default risk over a given set of clusters (e.g., sectors). The comparison between these two distributions enables us to monitor and quantify a counter-cyclical Recession Risk component separately from the "standard" default risk measure, as shown in Sections 2.5 and 2.6. Figure 1 displays a diagram summarizing the steps involved in the model calibration and application, as described in the remainder of this section.

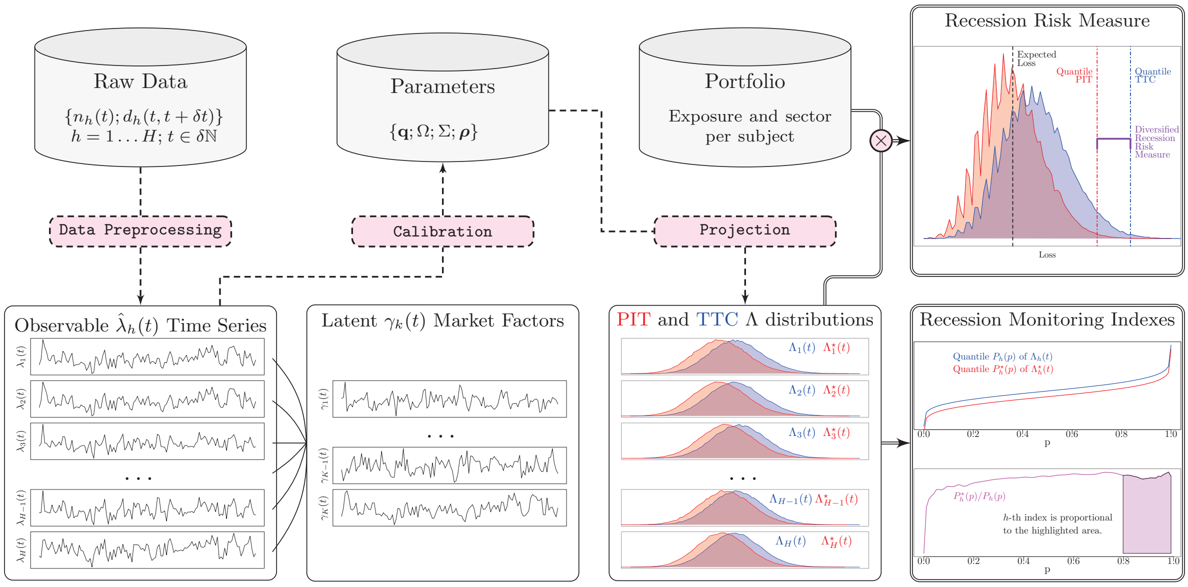

Figure 1. Schematics of the RecessionRisk$^+$ model's calibration and application.

Remark 2. *The formal scope of the RecessionRisk$^+$ definition goes as follows. Assumptions 3 and 4 are foundational. Proposition 1 proves that the model is well-posed and enables the construction of a practical form for the TTC hazard rates $\Lambda_h(t)$ $(h = 1 \ldots H)$ in Equation* (15)*. Proposition 2 provides the corresponding form for the PIT hazard rates $\Lambda^\star_h(t)$ in Equation* (17)*. Proposition 3 enables the model's parameterization at both the considered time scales and, together with the remainder of Section 2.4, constructs a complete calibration procedure. The framework allows to construct objects to quantify Recession Risk as a function of $\boldsymbol{\Lambda}$ and $\boldsymbol{\Lambda}^\star$ distributions. Indeed, the model is made applicable by the introduction of some procyclical quantifiers of Recession Risk, such as $\overline{RR}_{hp}(t)$ in Equation* (31) *(see Section 2.5), and also a couple of countercyclical quantifiers U_R, S_R (see Section 2.6)—all of them being applied to real case studies in the following Section 3. There is no claim of completeness concerning the set of quantifiers designed in this work, as the model user can define other quantifiers in the same RecessionRisk$^+$ framework depending on the specific application's requirements.*

2.4. Model Calibration

The model introduced in Section 2.3 defines the loss distribution over the time horizon $(t_0, t_0 + \Delta]$ both directly (i.e., uniperiodal, unconditioned loss distribution) and in a roll-over framework (i.e., m subsequent projection steps onto δ-long periods, each of them conditioned to the former). To fully apply the model, we need to calibrate all the parameters introduced above. Indeed, the parameters set

$$\{\mathbf{q}; \sigma^2; \Omega\} \in [0,1]^H \times \mathbb{R}^K_+ \times [0,1]^{HK},$$

estimated in t_0, is needed to evaluate in t_0 the loss distribution at $t_0 + \Delta$ (i.e., onto the long time scale directly). For the sake of convenience, in the following we consider the factor loadings' matrix $\Omega \in [0,1]^{HK}$ without the row of idiosincratic terms ω_{h0} ($h = 1 \ldots H$), as it is fully specified by the normalization constraints in Equation (3). In addition, the parameters

$$\{\boldsymbol{\rho}; \boldsymbol{\gamma}(t-\delta)\} \in [0,1]^K \times \mathbb{R}_+^K$$

must also be specified to evaluate in t the loss distribution at $t+\delta$ (i.e., first step of the roll-over process onto the short time scale).

First, let us introduce two observable counters that are needed to build the estimators for the quantities listed above: $n_h(t) := |c_h(t)|$ is the number of solvent debtors belonging to the h-th cluster c_h at time t, and $d_h(t, t+\delta) := \sum_{i \in c_h} \tilde{Y}_i$ is the number of debtors defaulted onto the interval $(t, t+\delta]$ and belonging to the same cluster c_h. For the sake of brevity, the time specifications will be omitted unless needed.

From the counters n_h and d_h, we can define $f_h(t) := d_h(t, t+\delta)/n_h(t)$, that is the empirical frequency of default measured on the interval $(t, t+\delta]$. Given that f_h estimates the Bernoulli distribution parameter $\tilde{p}_h$, we can introduce the estimator

$$\hat{\lambda}_h := -\ln(1 - f_h) \tag{20}$$

As $(1-f_h)n_h \sim \text{Binomial}(1-\tilde{p}_h, n_h)$, the standard error associated with f_h scales as $n_h^{-1/2}$. Assuming large clusters $n_h \gg 1$ ($h = 1 \ldots H$), the error is negligible. Since Equation (14) and Assumption 3 imply that $q_h = m\mathbf{E}[\lambda_h]$, we can estimate q_h as

$$\hat{q}_h := m\langle \lambda_h \rangle = \frac{m}{N} \sum_{j=1}^{N} \hat{\lambda}_h(j\delta). \tag{21}$$

It is possible estimating q_h based on the long-time-scale observables as well, because Equation (15) implies that $m\mathbf{E}[\lambda_h] = \mathbf{E}[\Lambda_h]$.

Also the parameters $\{\sigma^2; \Omega\}$ can be estimated based on both the considered time scales, through the result reported below.

Proposition 3. *Assumption 3 implies the following*

$$\boldsymbol{cov}[\boldsymbol{\Lambda}, \boldsymbol{\Lambda}] = m\,\boldsymbol{cov}[\boldsymbol{\lambda}, \boldsymbol{\lambda}] + 2\sum_{x=1}^{m-1} (m-x)\boldsymbol{cov}[\boldsymbol{\lambda}, {}_x\boldsymbol{\lambda}] = Q\Omega^T \Sigma \Omega Q, \tag{22}$$

where $\boldsymbol{\Lambda} := (\Lambda_1, \ldots, \Lambda_H)$, $\boldsymbol{\lambda} := (\lambda_1, \ldots, \lambda_H)$, ${}_x\boldsymbol{\lambda}(t) := \boldsymbol{\lambda}(t + x\delta)$, $\Sigma := \boldsymbol{diag}[\boldsymbol{\sigma}^2]$, *and* $Q := \boldsymbol{diag}[\boldsymbol{q}]$.

Proof. From Equations (14) and (15) it follows

$$\begin{aligned}
\mathbf{cov}[\Lambda_h, \Lambda_{h'}] &= \mathbf{cov}\left[\sum_{j=1}^{m} \lambda_{hj}, \sum_{j'=1}^{m} \lambda_{h'j'}\right] \\
&= \sum_{j=1}^{m} \mathbf{cov}\left[\lambda_{hj}, \lambda_{h'j}\right] + 2\sum_{j=1}^{m-1} \sum_{j'=j+1}^{m} \mathbf{cov}\left[\lambda_{hj}, \lambda_{h'j'}\right] \\
&= m\,\mathbf{cov}[\lambda_h, \lambda_{h'}] + 2\sum_{x=1}^{m-1} (m-x)\mathbf{cov}[\lambda_h, {}_x\lambda_{h'}],
\end{aligned}$$

where, for the sake of brevity, we have adopted the short notation $\lambda_h(t + j\delta) = \lambda_{hj}$ and omitted the time argument unless needed. Since Assumption 3 implies

$$\mathbf{cov}[\lambda_h, {}_x\lambda_{h'}] = \frac{1}{m} q_h q_{h'} \sum_k \varrho_{xk} \omega_{hk} \omega_{h'k} \xi_k^2 \sigma_k^2,$$

then we have

$$\begin{aligned}\mathbf{cov}[\Lambda_h, \Lambda_{h'}] &= q_h q_{h'} \sum_k \left[1 + 2\sum_{x=1}^{m-1} \varrho_{xk}\left(1 - \frac{x}{m}\right)\right] \omega_{hk}\omega_{h'k}\xi_k^2\sigma_k^2 \\ &= q_h q_{h'} \sum_k \omega_{hk}\omega_{h'k}\sigma_k^2,\end{aligned}$$

where the simplification above is due to the definition of ξ_k. □

As detailed in [22], the short-time-scale estimator is more convenient in terms of lesser error. Thus, we choose to calibrate the parameters $\{\sigma^2; \Omega\}$ by applying the SNMF decomposition technique (see [21,22]) to the matrix $\hat{A}$, whose elements are defined as follows:

$$\hat{A}_{hh'} := \frac{1}{\hat{q}_h \hat{q}_{h'}} \left(m\, \mathbf{c\hat{o}v}\left[\hat{\lambda}_h, \hat{\lambda}_{h'}\right] + 2\sum_{x=1}^{m-1} (m - x)\mathbf{c\hat{o}v}\left[\hat{\lambda}_h, {}_x\hat{\lambda}_{h'}\right] \right), \tag{23}$$

where $\mathbf{c\hat{o}v}[\cdot]$ is the sampling estimator of covariance.

Given $\hat{\Omega}$ and $\hat{\Sigma}$ estimated by decomposing $\hat{A}$, we can calibrate the ACF $\boldsymbol{\varrho_k}$ $(k = 1 \ldots K)$ by the following relation between observable quantities and parameters, once again directly implied from Assumption 3:

$$\frac{1}{m}\mathbf{var}[\lambda_h] = \sum_k \omega_{hk}^2 \xi_k^2 \sigma_k^2.$$

Hence, the ACF parameters $\boldsymbol{\rho} \in [0, 1]^K$ introduced in Assumption 4 are estimated as the solution of the following optimization problem:

$$\hat{\boldsymbol{\rho}} = \mathbf{argmin}_{\boldsymbol{\rho}} \left|\left| \mathbf{diag}\left[\hat{\Omega}^T \mathbf{diag}\left[\hat{\sigma}^2 \xi^2(\boldsymbol{\rho})\right] \hat{\Omega}\right] - \frac{1}{m}\mathbf{v\hat{a}r}\left[\hat{\lambda}\right] \right|\right|, \tag{24}$$

where $\mathbf{v\hat{a}r}[\cdot]$ is the sampling estimator of variance.

Remark 3. *The optimization problem stated in Equation* (24) *can still be utilized to calibrate the ACF, even if Assumption 4 is modified to choose a functional form of* ϱ_{xk} *other than* $\rho_k^{|x|}$.

Finally, $\boldsymbol{\gamma}(t - \delta)$ can be inferred by as the solution of the problem

$$\hat{\boldsymbol{\gamma}} = \mathbf{argmin}_{\boldsymbol{\gamma}} \left|\left| \frac{\hat{q}_h}{m} \left(\hat{\omega}_{h0} + \sum_{k=1}^{K} \hat{\omega}_{hk}\, \gamma_k \right) - \hat{\lambda}_h(t - \delta) \right|\right|. \tag{25}$$

Section 3.1 reports a numerical validation of the calibration procedure discussed above.

2.5. Application as a Recession Risk Forecasting Index

Section 2.3 introduces two random variables $\boldsymbol{\Lambda}, \boldsymbol{\Lambda}^\star \in \mathbb{R}^H$, representing the unconditioned and conditioned one-year hazard rates per cluster $h = 1, \ldots, H$, respectively. Section 2.4 describes the calibration procedure needed to specify their distributions fully.

In the remainder of Section 2, we describe the two applications of the framework depicted so far. First, let us introduce the random variables needed in this subsection and the following one. As anticipated above, $\boldsymbol{\Lambda}^\star(t, \overline{\boldsymbol{\gamma}})$ can be interpreted as the array of one-year PIT hazard rates. Thus, the one-year PIT default probability $P_h^\star$, default event $y_i^\star$, and default frequency $F_h^\star$ of the debtors belonging to the h-th cluster can be written as

$$P_h^\star := 1 - \exp(-\Lambda_h^\star), \tag{26}$$

$$y_i^\star \sim \text{Bernoulli}(P_h^\star), \quad i \in c_h, \tag{27}$$

$$F_h^\star := \frac{\tilde{d}_h^\star}{n_h}, \quad \tilde{d}_h^\star := \sum_{i \in c_h} y_i^\star \sim \text{Binomial}(P_h^\star, n_h), \tag{28}$$

while the replacement $\Lambda_h^\star \to \Lambda_h$ provides the corresponding TTC random variables: P_h, y_i, and F_h. It is worth recalling that our model—like $CR+$—is doubly stochastic, as $P_h^\star, P_h$ are also random variables. Since P_h is a TTC variable, its right-tail values are expected to reproduce the frequency-of-default level observed during past recession crises, provided that the time series $\hat{f}_h$ considered in calibration is deep enough in the past to be considered TTC.

Assuming that the model is calibrated based on adequate time series, we determine the next year $(t, t+1]$ "recession" scenario of default probability for the h-th cluster as $\overline{P}_{hp}(t) := \Phi_h^{-1}(p)$, where $\Phi_h(\cdot)$ is the cdf of $P_h(t)$ and $p \in (0,1)$ is a given tolerance level.

Loosely speaking, to understand how far c_h is from experiencing a recession in a p-worst-case scenario, we need to compare $\overline{P}_{hp}$ with the corresponding PIT quantity $\overline{P}_{hp}^\star$. Thus, two plain and readable definitions of a Recession Risk index per cluster are

$$RR_{hp}(t) := \frac{\overline{P}_{hp}^\star(t)}{\overline{P}_{hp}(t)} \in \mathbb{R}^+, \tag{29}$$

$$RR'_{hp}(t) := \frac{\min\{\overline{P}_{hp}^\star(t); \overline{P}_{hp}(t)\}}{\overline{P}_{hp}(t)} \in [0,1]. \tag{30}$$

As $RR_{hp}, RR'_{hp} \simeq 1$, the tail values of the PIT default probability distribution are closer to the corresponding TTC values; thus, recession events similar to the ones observed in the past are more likely. On the other side, $RR_{hp}, RR'_{hp} \ll 1$ implies that the tail of the current PIT distribution estimated in t is not comparable with the corresponding TTC distribution and, thus, the h-th cluster is far from experiencing a recession phase.

The "min" term introduced in the RR'_{hp} definition bounds the index domain to the compact interval $[0,1]$. Hence, RR'_{hp} is dedicated only to understand whether the h-th cluster is far from experiencing a recession phase or not. Without the "min" term, it is not forbidden that the currently observed PIT distribution $\Phi_h^\star$ exhibits a "wider" right tail that the TTC distribution Φ_h. This may be the case when approaching a recession worse than all previously observed in the time series available for calibration purposes. In such a case, RR_{hp} may provide an insight concerning the magnitude of the possible recession phase that is approaching or occurring.

The index definition is anchored to the choice of p. At an increasing p, we are comparing less likely and more extreme scenarios of the PIT and the TTC distributions. Hence, $p \lessapprox 1$ should be chosen if we want to have early warnings of rare, extreme recession periods. On the other hand, a lesser p value—such as the 3rd quartile $p = 0.75$—should be chosen when aiming to monitor more frequent but less intense periods of simultaneous default events.

A comprehensive comparison between the PIT and TTC hazard rate tails is also feasible by considering an averaged index, such as

$$\overline{RR}_{hp}(t) := \frac{1}{1-p}\int_p^1 RR_{hp'}(t)dp' \in \mathbb{R}^+, \tag{31}$$

$$\overline{RR}'_{hp}(t) := \frac{1}{1-p}\int_p^1 RR'_{hp'}(t)dp' \in [0,1]. \tag{32}$$

Indexes proposed in Equations (29)–(32) provide nontrivial information in at least two aspects. First, they compare the right tails of two forward-looking probability distributions arising from the proposed model. Second, each cluster's index is evaluated based on the dependency structure among the clusters. Hence, the indexes can highlight a potential crisis situation even when a given cluster's currently observed default frequency is under the cluster's average level.

A numerical application of the proposed indexes to a real-world historical dataset is provided below in Section 3.2.

2.6. Application as a Solvency II Recession Risk Internal Model

In the following, RecessionRisk^{+} is applied as a PIM in the Solvency II framework to jointly measure C&S Premium Risk and Recession Risk.

First, let us introduce the random variable $L^{\star}$, which is the PIT one-year loss distribution generated by a given C&S portfolio.

$$L^{\star} := \sum_{i} \ell_i y_i^{\star}, \tag{33}$$

where ℓ_i is the loss-given default generated by the i-th debtor. ℓ_i can be either deterministic or stochastic, with $L^{\star}$ being easily simulated in both cases. When dealing with portfolios that are homogenous at cluster or higher level, $L^{\star}$ can be simplified as $L^{\star} = \sum_h \ell_h d_h^{\star}$.

The Premium Risk SCR S_P represents the unexpected one-year loss level at the 99.5% quantile for a given Line of Business. Thus, we can quantify the C&S Premium Risk measure U_P as a risk capital evaluated on the $L^{\star}$ distribution.

$$U_P = \Phi_{L^{\star}}^{-1}(0.995) - \mathbf{E}[L^{\star}], \tag{34}$$

where $\Phi_{L^{\star}}(\cdot)$ is the cdf of $L^{\star}$. As discussed above in Section 2.5 with regard to each considered cluster, the C&S portfolio is likely to experience a recession phase when the right tail of $\Phi_{L^{\star}}$ is approaching the corresponding TTC distribution Φ_L. Hence, our framework provides a "natural" dampening measure U_R to quantify the Recession Risk SCR:

$$U_R := \max\{0; \Phi_L^{-1}(0.995) - \Phi_{L^{\star}}^{-1}(0.995)\}. \tag{35}$$

Using the same concept underlying Equation (30), U_R widens as the "global" TTC loss distribution $\Phi_L(\cdot)$ exhibits a right tail that is wider than the one of the corresponding "local" PIT distribution $\Phi_{L^{\star}}(\cdot)$. Conversely, when $\Phi_{L^{\star}}(\cdot)$ tail loss scenarios are similar to or worse than the ones implied by $\Phi_L(\cdot)$, the capital requirement generated by the Recession Risk component lowers, releasing the company's Own Funds. The "max" term is introduced to prevent a negative capital requirement in extremely severe recession scenarios.

Intuitively, the measures U_P and U_R could be summated, resulting in a prudential measure of unexpected loss $\Phi_L^{-1}(0.995) - \mathbf{E}[L^{\star}]$, where the "max" term is not considered as being introduced only to cope with the Solvency II balance sheet legal requirements and with the expected one-year loss being evaluated based on the PIT loss $L^{\star}$, which provides a more reliable short-term estimate than the TTC loss L by definition.

Although reasonable, this linear aggregation mechanism between C&S Premium Risk and Recession Risk SCR components does not cope with the Solvency II Standard Formula's aggregation rules. Indeed, the Standard Formula requires that

$$\begin{aligned} S_U &= \sqrt{S_P^2 + 2\rho_{PC} S_P S_C + S_C^2}, \\ S_C &= \sqrt{S_D^2 + S_R^2}, \end{aligned} \tag{36}$$

where $\rho_{PC} = 25\%$, S_R is the Recession Risk SCR, S_D is the Catastrophe Default Risk SCR, S_C is the Catastrophe Human-made Risk SCR, and S_U is the Underwriting Risk SCR.

Equation (36) holds specifically under the assumption that all the risk components different from S_P, S_D, and S_R are null and that C&S is the only Line of Business to be considered at risk. Imposing $S_R = U_R$ and $S_P = U_P$ implies that $S_U < U_P + U_R$ because we exploit a diversification effect introduced in the Standard Formula aggregation structure that is not featured in our model structure.

To fix this, we can insert our PIM in the Standard Formula structure, imposing that

$$\begin{cases} S_U &= U_R + U_P, \\ S_P &= U_P, \\ S_D &= 0. \end{cases} \tag{37}$$

Remark 4. *Actually, $S_D \gneqq 0$ by construction. In fact, $S_D = lgd(LE_1 + LE_2)$, where LE_i ($i = 1, 2$) are the first and the second largest exposures in the considered C&S portfolio, and $lgd = 10\%$ is a loss-given-default coefficient fixed by EIOPA. Hence, any non-empty portfolio that exhibits $S_P, S_R > 0$ also has a positive S_D measure. Nonetheless, $S_D = 0$ copes with the other constraints listed in Equation* (37)*, implying that the measure $U_R + U_P$ is distributed only between S_P and S_R.*

Equations (36) and (37) are satisfied by the following equation:

$$\begin{aligned} S_R &= -\rho_{PC} U_P + \sqrt{(U_R + U_P)^2 - (1 - \rho_{PC}^2) U_P^2}, \\ S_P &= U_P. \end{aligned} \tag{38}$$

Equation (38) allows the application of RecessionRisk$^+$ to the Solvency II Recession Risk SCR without any undue diversification benefit. It is worth noting that S_R behaves as required. In fact, it holds $S_R \geq 0$ and $\partial_{U_R} S_R > 0$ by construction. Further, it is easily verified that $S_R > 0$ if $\Phi_{L^\star}^{-1}(0.995) < \Phi_L^{-1}(0.995)$.

Applying Equation (38) to the Premium Risk and Recession Risk measures, all the other SCR components can be evaluated and aggregated according to the Standard Formula specifications, including S_D.

We can heuristically assess the degree of prudentiality obtained with the map $U_R \to S_R$ by considering the model's behavior at the trough and the peak of the economic cycle separately.

At the trough of the cycle, it is expected that $U_R \ll U_P$. Considering the first-order MacLaurin series approximation of $S_R(U_R)$ in Equation (38), we have

$$S_R \approx \frac{1}{\rho_{PC}} U_R.$$

Conversely, at the peak of the cycle, it is expected $U_R \gg U_P$. Considering the first-order MacLaurin series approximation of S_R as a function of $\rho_{PC} U_P$, we have

$$S_R \approx U_R + (1 - \rho_{PC}) U_P.$$

Hence, U_R and S_R are comparable at the peak of the cycle, while S_R is expected to be about four times U_R during recession periods, neutralizing the undue diversification benefit and satisfying the prudential behavior required from a PIM in the Solvency II framework.

A numerical application of the proposed Partial Internal Model to a real-world historical dataset is provided below in Section 3.3.

3. Numerical Applications

This section presents a selection of numerical experiments and case studies concerning the theoretical framework presented so far.

Section 3.1 discusses a numerical experiment performed to verify the absence of bias in the estimators introduced in Section 2.4, performing the $RR+$ calibration over a set of simulated scenarios, where the "true" parameters behind the simulated default frequencies are a priori chosen and, thus, known.

Sections 3.2 and 3.3 describe two real-world case studies where the model applications introduced above in Sections 2.5 and 2.6 are implemented and tested.

3.1. Validation of the Proposed Calibration Techniques on a Simulated Dataset

Generally speaking, unbiasedness is a desirable property for the estimators of any model's parameters. To assess an estimator's unbiasedness analytically, one must derive its distribution function and compute its expected value. Numerically, a Monte Carlo simulation allows one to check whether the estimators show clear deviations from the hypothesis of unbiasedness.

We have investigated the (possible) bias of the estimators introduced in Section 2.4 with a dedicated Monte Carlo experiment, finding that the sample averages of the estimated values of the parameters are in agreement, within confidence interval implied from Monte Carlo error, with the values used—in input—to generate the Monte Carlo samples mimicking real datasets.

The experiment was carried out as follows: We have considered an "infinite size" portfolio composed of $H = 8$ clusters and a market described by $K = 3$ sectors. The "infinite size" assumption corresponds to assuming that the size $n_h(t)$ of each h-th cluster population is large enough that the contribution of the binomial error to the estimate of $p_h(\mathbf{\Gamma})$ is negligible. The numerical findings of [22] justify this assumption and allow to speed up the test.

Given the portfolio, we have generated $n_S = 2.000$ trajectories of the state variable $\boldsymbol{\gamma} = (\gamma_1, \ldots, \gamma_K)$, taking into account autocorrelation in the trajectories. The trajectories start at time $t_0 = 0$ and end at time $t_n = 10$ years, with a monthly time step ($m = 12$). In each trajectory, autocorrelation was produced by generating, for each $k \in [1, K]$, the vector $(\gamma_k(t_1), \gamma_k(t_2), \ldots, \gamma_k(t_n))$ from a multivariate distribution having a dependence structure described by a normal copula function with correlation matrix $\Sigma_k^{(\gamma)}$, and $\text{Gamma}\left(\sigma_k^{-2}\,\xi_k^{-2}\,m^{-1}, \sigma_k^2\,\xi_k^2\,m\right)$ marginal distributions. Following Assumption 4, the elements of the K correlation matrices were chosen as $(\Sigma_k^{(\gamma)})_{ij} = \rho_k^{|i-j|}$.

Given the state variable trajectories, for each trajectory and each cluster in the portfolio, we computed the default probabilities $p_h(t, \boldsymbol{\gamma}(t))$ ($h = 1, \ldots, H$). In this way, we obtained a set of n_S datasets, each mimicking a possible time series of the realized default frequencies $f_h(t)$ under the infinite size assumption mentioned above.

The next step is the estimate of the $(H \times H)$ $\hat{A}$ matrix, as defined in Equation (23). Possibly, this operation implies a regularization of the estimated covariance matrix of the hazard rates λ_h, which is performed using a standard algorithm [38]. We have checked that the inferred $\hat{A}$ matrix per scenario reproduced well to the "true" A matrix utilized to simulate the scenarios.

The following step is the decomposition of $\hat{A}$ in $\widehat{\Omega}$ and $\widehat{\Sigma}$ using the SNMF, which is performed as in [21,22]. Numerically, this is the most delicate phase of the calibration since the solution to the numerical problem is not unique; in fact, different approximate solutions exist and are equally acceptable. Possibly, the factorization can use a regularization technique to identify the solution closest to an a priori solution. However, in this work, no regularization was introduced. Thus, we tested the unbiasedness of the estimate on the elements of the diagonal of the matrix $\Sigma_\Lambda := \mathbf{cov}[\mathbf{\Lambda}, \mathbf{\Lambda}]$ rather than on the elements of the matrices $\widehat{\Omega}$ and $\widehat{\Sigma}$ separately. Since Σ_Λ depends on the product of Σ and Ω, this approach is a test that the factorization respects all constraints.

Figure 2 shows the distributions (as boxplots) of the estimates obtained in the n_S scenarios versus the input values (as red dots). Since the average values of the boxplots cope with the input values, the possible unknown bias is deemed sufficiently small to be neglected.

After matrix factorization is performed, the values of ρ_k ($k = 1, \ldots, K$) parameters can be inferred using Equation (24).

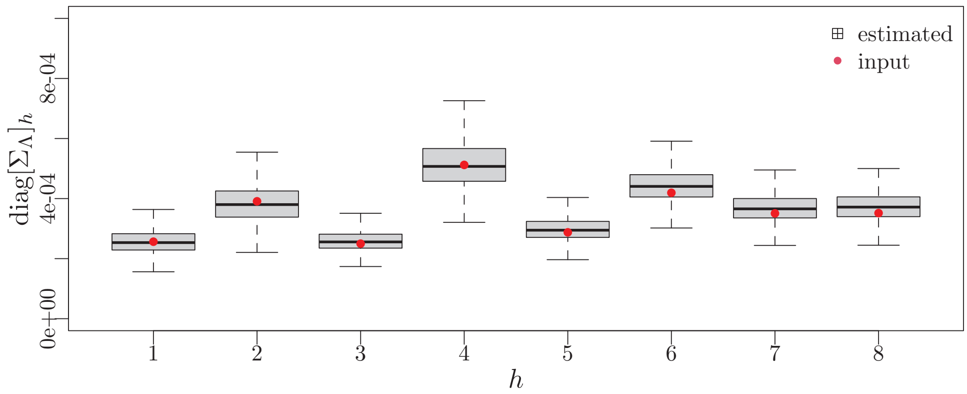

Figure 2. Comparison of the estimated values' distribution of the diagonal elements of Σ_Λ (shown as boxplots) and the corresponding input values to the simulated trajectories (red points), for each of the $H = 8$ clusters.

The last part of the numerical exercise is the reproduction of $\lambda_h(t)$ observed at the end of the time series. This computation involves all the estimated parameters and the inference of $\hat{\gamma}(t - \delta)$, which can be estimated as the solution to the problem stated in Equation (25). Figure 3 shows each cluster's distribution of the $\lambda_h(t)$ relative error. The distributions are nicely centered on zero, which provides evidence in favor of the estimator's unbiasedness.

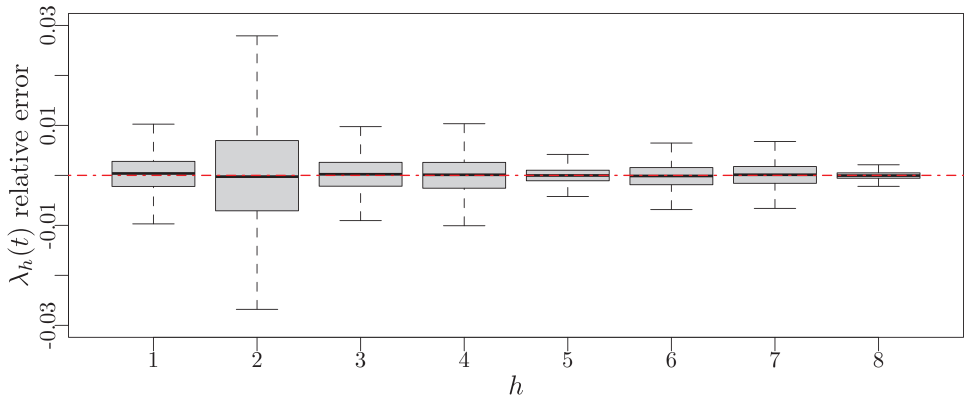

Figure 3. Distributions (boxplots) of the relative error in the estimate of $\lambda_h(t)$ at the end of the time series for each of the $H = 8$ clusters. The red dashed line is the zero relative error level.

Given that the chosen numerical scenarios were generated based on realistic parameters, we are confident that the numerical procedure used in the estimate is robust enough both from the theoretical and practical points of view.

3.2. The 2012 Italian Sovereign Debt Crisis: A RR+ Indexes' Case Study

In the 2011–2012 period, a complex crisis concerning the spread of sovereign debt arose across Europe. In Italy, significant effects on the real economy were apparent starting from 2012, following the approval of a large austerity plan in September 2011 [39]. The Italian parliament approved such a plan to contrast the crisis and its effect on the spread dynamics of Italian obligations. Discussing the causes and the evolution of that recession period is beyond the purposes of this work; however, the literature offers deep and accurate analyses that investigate those years under various perspectives (see, e.g., refs. [40,41] and the references therein).

Nonetheless, it is worth studying the behavior of $RR+$ through the time series of the Italian insolvency rates before and after the 2012 crisis. Indeed, these are publicly available from the Italian banking system and offer an ideal test bed to our model, as it can be calibrated based on observations preceding the crisis and then applied to define indexes to be observed during the crisis, verifying its ability to detect the different regimes occurring after 2012. As detailed in Section 2.4, the only dataset needed to calibrate $RR+$ is the couple of counters $n_h(t)$ and $d_h(t)$, per cluster and past the observation date. Indeed, their ratios $f_h(t)$ (like the Italian insolvency rates considered in this application) enable the full parameterization of the model. No other financial or economic information is needed.

The data utilized to calibrate the $RR+$ model for this application are the historical time series of bad loan rates supplied by the Bank of Italy. "Bad loans" are a subcategory of the broader class "Non-Performing Loans", and they are defined as "exposures to debtors that are insolvent or in substantially similar circumstances" [42]. In particular, the chosen dataset is composed of the quarterly historical series TRI30496 ($m = 4$, choosing 1 year as the long time scale) over 14 years (from 1 January 2004 to 31 December 2017). The time series are publicly available at [43] and are supplied by six distinct customer sectors ($H = 6$), listed in Table 1.

Table 1. Economic sectors available in the TRI30496 historical dataset.

h	Sector's Id	Description
1	600	Consumer households
2	S11	Non-financial companies
3	S12BI7	Financial companies other than monetary financial institutions
4	S13	General government
5	S14BI4	Producer households
6	S15BI1	Limited and general partnerships, informal partnerships, and de facto companies

As explained above, this application of $RR+$ focuses on the model's ability to detect—or even to anticipate—the effect of the 2011–2012 crisis on the default rates observed in the following years, being calibrated to time series up to the end of 2010. Figure 4 displays the model performance across the available sectors. The index $\overline{RR}_{hp}(t)$ introduced in Equation (31) is compared with the default rate measured over the following period $(t, t + 1 \text{ year}]$. The only parameter updated on a quarterly basis is the array $\overline{\gamma}(t - \delta)$, representing the last inferred value of the short-scale latent factors, based on the information available at t. All the other parameters are fixed at the 31 December 2010 filtration, estimated following the techniques described in Section 2.4. Re-performing a quarterly complete calibration would have been feasible, and the model would have a fair advantage from this. Nonetheless, calibrating the model only once highlights its robustness in a real-world context. Figure 4 shows that the index dynamics mimic or anticipate the observed default rate dynamics. Although noise and local deviations of the index from a posteriori observations are appreciable, the result is even more remarkable considering that $RR+$ is a reduced-form autoregressive model and does not consider any exogenous information on the macroeconomic or financial context at all.

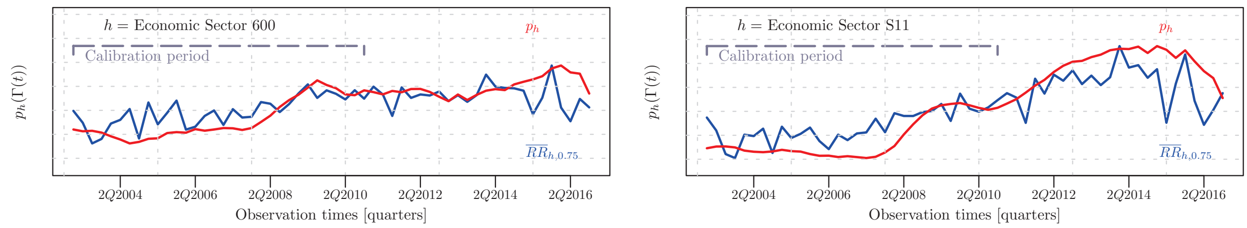

Figure 4. *Cont.*

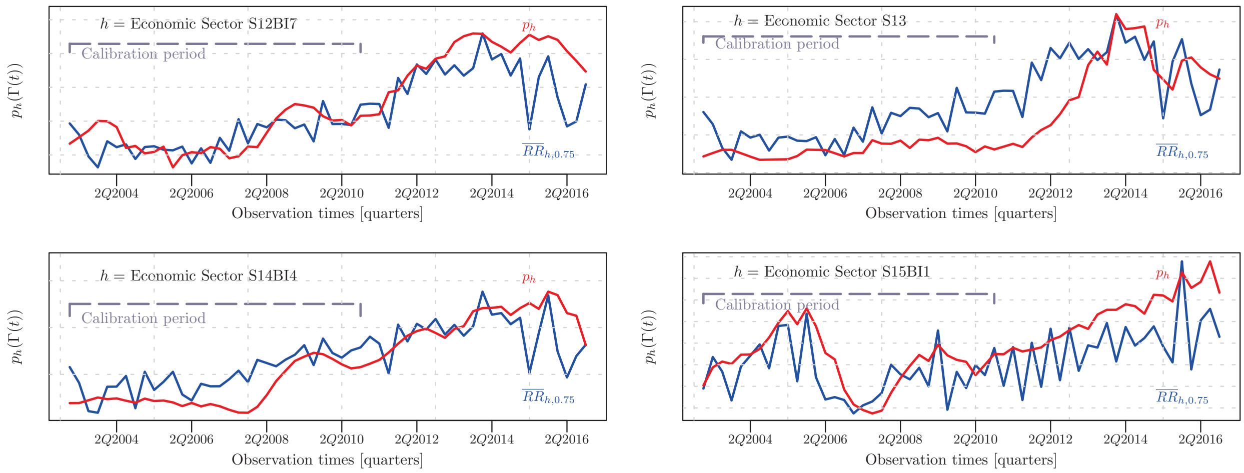

Figure 4. Index $\overline{RR}_{hp}(t)$ ($p = 0.75$, blue line) scaled and compared to the a posteriori observed 1-year default frequency time series (red line) for each sector listed in Table 1.

3.3. Numerical Comparison Between the RR+ Internal Model and the Standard Formula

In Section 2.6, the $RR+$ model is utilized to quantify the Recession Risk S_R and the Premium Risk S_P components of the Solvency Capital Requirement (SCR) for an insurance company operating in the Solvency II's Segment 6 (credit and suretyship insurance). Further, the same framework also provides a measure of the expected loss $\mathbf{E}[L^\star]$, which represents the Best Estimate Liabilities component of the Premium Provision for the same segment. We now introduce some simplifying assumptions in order to build a toy Solvency II balance sheet with the available elements and compare its behavior when computed through the Standard Formula and $RR+$ internal model. First, we assume claims are being paid instantly (i.e., no reserve provisions and no reserve risk) and that lapse events do not apply to the policies at risk in the company's portfolio. The latter assumption implies the absence of Lapse Risk, which is realistic for a Segment 6 portfolio. The policies are assumed to expire precisely after one year, which, once again, is a realistic assumption for the considered line of business and allows us to exclude the future premiums contribution from our Standard Formula determination of the SCR.

Further, the Risk Margin RM is assumed to be proportional to the SCR, enabling the computation of technical provisions from the available elements. Conversely, the policy premiums are assumed to be proportional to $\mathbf{E}[L^\star]$, following a classic approach (see, e.g., ref. [13] and the references therein). This assumption enables the computation of the volume measure needed for the Standard Formula Premium Risk and Catastrophe Recession Risk. The other components of the SCR not directly related to the insurance policies are neglected (e.g., operational risk; market, credit, and counterparty risks; deferred taxes and intangible assets). Finally, Catastrophe Default Risk is kept constant through time, assuming that the two largest exposures in the portfolio remain the same two debtors or contractors.

The (toy) Solvency II liabilities described above can be fully quantified with the $RR+$ model and the Standard Formula. We considered the dataset described in Section 3.3 as realistic quarterly dynamics of default probabilities. Each of our simulated debtors and contractors has an exposure inversely proportional to their default probabilities to represent the greater risk appetite of an insurer for the less risky subjects, and they are distributed among all the six economic sectors reported in Table 1. Instead of a uniform distribution, we have chosen to distribute them, introducing a negative dependency between n_h and q_h. Doing so, we have a realistic portfolio with "many" small exposures on subjects characterized by low creditworthiness and, conversely, a few exposures on high-standing subjects.

Figure 5 shows the behavior of the Solvency II liabilities across the quarters if the $RR+$ model is applied. The total liabilities level remains remarkably stable through time, although nontrivial dynamics of the underlying components are observed. This fact shows that the dampening mechanism originally conceived by EIOPA is restored by applying our model. Conversely, Figure 6 displays the effect of the Solvency 2 Standard Formula quantification of the same liabilities. Given the simplifications adopted in the Standard Formula's SCR component for Recession Risk, the dampening mechanism is absent, and the level of total liabilities is subject to evident variations through time.

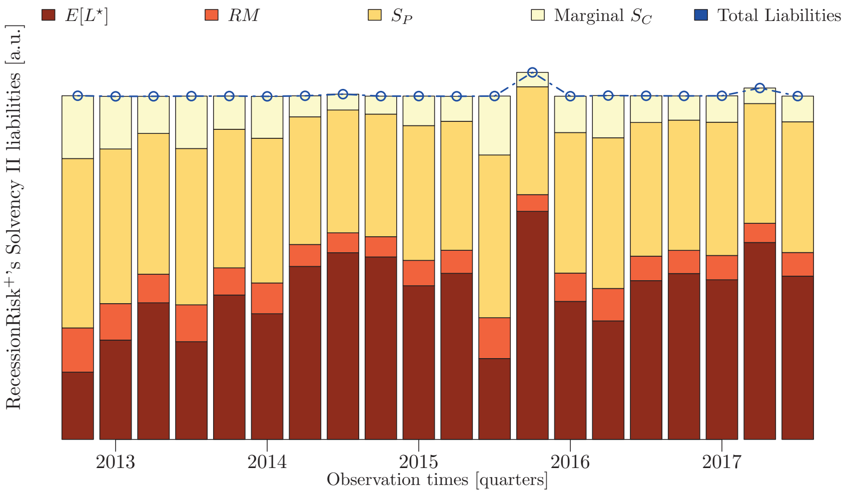

Figure 5. Solvency II liabilities applying the $RR+$ model (quarterly simulations). The liabilities' level is displayed on a logarithmic scale and with arbitrary units of economic value.

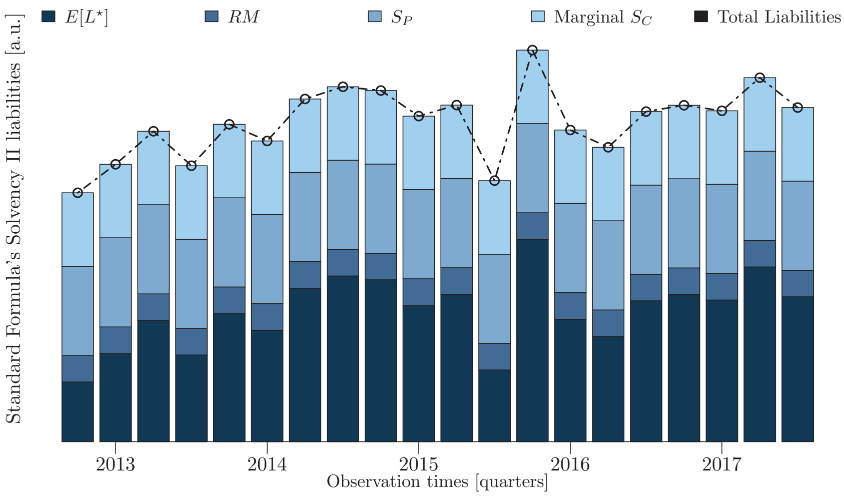

Figure 6. Solvency II liabilities applying the Solvency II Standard Formula (quarterly simulations). The liabilities' level is displayed on a logarithmic scale and with arbitrary units of economic value.

4. Conclusions

In this work, we have proposed a new model to measure Recession Risk based on the collective dynamics of default events generated by a population that can be clustered in distinct, internally homogenous sectors. The RecessionRisk$^+$ model can be regarded as a reduced-form latent-factors' model, like the classic CreditRisk$^+$ from whom it was inspired. Nonetheless, there are relevant differences between the two models. The Bernoulli representation of default events was restored (instead of the Poisson approximation adopted by $CR+$ and moving the copula application from default probabilities to hazard rates), and the latent market factors were defined over multiple time scales, enabling a roll-over framework both in calibration (past observation times) and in projection (future observation times). Further, each market factor was autocorrelated, introducing a model memory that allowed different conditional distributions of future default probabilities to interfere depending on the recently observed default frequencies.

The RecessionRisk$^+$ model has practical applications that make it relevant in real-world scenarios. It enables the estimation of the likelihood of a recession occurring in the next period, allowing for a comparison of multivariate through-the-cycle probability distribution and the corresponding point-in-time distribution within a consistent framework. This comparison facilitates the definition of a Solvency II internal model for Recession Risk that restores the dampening mechanism originally conceived by EIOPA, as we have numerically verified. The model can also be applied to define a set of non-macroeconomic indexes that allow for the separate monitoring of an economy's proximity to a recession period per economic sector. Numerical simulations show that the model's performance meets the a priori expectations in both the proposed applications.

A limitation of this research that may be addressed in further works is that a part of the analytical results is based on a "low-volatility assumption" that enables the approximation of the Gamma distributions used to describe the market factors as normal distributions. Further analytical developments may generalize the closed-form results provided by this work, weakening or even removing that approximation, although the model can be utilized in its present version without any practical limitations. Another possible analytical development is the investigation of the model's consistency, given an ACF functional form that is different from the one utilized in this work. The robustness of the currently proposed set of assumptions can be further investigated as well, both by measuring the model's performances on simulated data under different conditions and by applying the model to other historical datasets featuring peaks of default frequencies, such as the credit risk dynamics observed during the COVID-19 pandemic or the commodities prices crisis originated by the Russo–Ukrainian conflict.

This work also offers the possibility of relevant further developments from an applied perspective. Since the model is designed to describe the collective dynamics of absorbing events in a short-term-memory system, a change of semantics could enable applications of the model in other fields, such as epidemiology, where it is worth investigating the possible application of this model to measure the pandemic risk of a population, or herd behaviors when facing binary decisions, with particular—but not exclusive—reverence to the context of behavioral finance.

Author Contributions: Conceptualization, J.G. and L.P.; Methodology, J.G. and L.P.; Software, J.G.; Validation, J.G. and L.P.; Formal analysis, J.G. and L.P.; Investigation, J.G.; Writing—original draft, J.G.; Writing—review & editing, L.P. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: The data that support the findings of this study are openly available in "Banca d'Italia—Base Dati Statistica" [43].

Conflicts of Interest: All authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Appendix A. Auxiliary Results

In this section, we report some results needed in Section 2.3.

Proposition A1 (Gaussian limit)**.** *Let X be a Gamma-distributed random variable, such that*

$$X \sim \mathrm{Gamma}\left(\sigma^{-2}, \sigma^2\right).$$

In the limit $\sigma \to 0^+$, it holds that

$$X \sim \mathcal{N}(1, \sigma^2).$$

Proof. The X probability densities $dF(x)$ can be written as

$$dF(x) \propto x^{\sigma^{-2}-1} \exp\left(-\sigma^{-2}x\right)dx \tag{A1}$$

Since $\sigma^{-2} - 1 \xrightarrow{\sigma \to 0^+} \sigma^{-2}$ holds, we have

$$\lim_{\sigma \to 0^+} dF(x) \propto \exp\left(\frac{\ln x - x}{\sigma^2}\right)dx. \tag{A2}$$

By introducing the auxiliary variable $x' := x - 1$ and replacing $\ln(1 + x')$ with the first three terms of its MacLaurin series, Relation (A2) can be equivalently written as

$$\lim_{\sigma \to 0^+} dF_k\big(x(x')\big) \propto \exp\left(-\frac{x'^2}{2\sigma^2}\right)dx' \tag{A3}$$

Changing back $x' \mapsto x(x')$ in Equation (A3) completes the proof. □

Proposition A2 (Correlated Gaussian variables). *Let (X, Y) be a normally distributed random variable in $\mathbb{R}^2$, such that*

$$X, Y \sim \mathcal{N}\left(\mu = \begin{bmatrix} \mu_x \\ \mu_y \end{bmatrix}, \Sigma = \begin{bmatrix} \sigma_x^2 & \rho\sigma_x\sigma_y \\ \rho\sigma_x\sigma_y & \sigma_y^2 \end{bmatrix}\right).$$

Moreover, the conditional distribution $X|Y$ is normal, with

$$\begin{aligned} \mathbf{E}[X|Y = y] &= \mu_x + \rho\sigma_x \frac{y - \mu_y}{\sigma_y}, \\ \mathbf{var}[X|Y = y] &= \sigma_x^2(1 - \rho^2). \end{aligned}$$

Proof. Let us consider first the standardized bivariate variable

$$X', Y' := \frac{X - \mu_x}{\sigma_x}, \frac{Y - \mu_y}{\sigma_y}$$

The (X', Y') joint density function is

$$f(x, y) = \frac{1}{2\pi\sqrt{1 - \rho^2}} \exp\left[-\frac{x^2 - 2\rho xy + y^2}{2(1 - \rho^2)}\right],$$

and the Y' marginal density function is

$$f(y) = \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{y^2}{2}\right]$$

Thus, the $X'|Y'$conditional density function can be written as

$$f(x|y) = \frac{f(x, y)}{f(y)} = \frac{1}{\sqrt{2\pi(1 - \rho^2)}} \exp\left[-\frac{(x - \rho y)^2}{2(1 - \rho^2)}\right],$$

implying $\mathbf{E}[X'|Y' = y'] = \rho y'$ and $\mathbf{var}[X'|Y' = y'] = 1 - \rho^2$. By changing back $X', Y' \mapsto X, Y$, we have

$$\begin{aligned} \mathbf{E}[X = x(x')|Y' = y'(y)] &= \mathbf{E}\left[X = \mu_x + \sigma_x x' \middle| Y' = \frac{y - \mu_y}{\sigma_y}\right] \\ &= \mu_x + \sigma_x \mathbf{E}\left[X' = x' \middle| Y' = \frac{y - \mu_y}{\sigma_y}\right] \end{aligned}$$

and

$$\begin{aligned}\mathbf{var}[X = x(x')|Y' = y'(y)] &= \mathbf{E}\left[X = \mu_x + \sigma_x x' \middle| Y' = \frac{y-\mu_y}{\sigma_y}\right] \\ &= \sigma_x^2 \mathbf{var}\left[X' = x' \middle| Y' = \frac{y-\mu_y}{\sigma_y}\right]\end{aligned}$$

completing the proof. □

Appendix B. List of Acronyms

Table A1 below reports all the acronyms utilized in this work for better readability.

Table A1. Acronyms utilized in Sections 1–4.

Acronym	Meaning
ACF	Autocorrelation Function
C&S	Credit and suretyship insurance
CR+	CreditRisk$^+$
CSFB	Credit Suisse First, Boston
EAD	Exposure at default
EIOPA	European Insurance and Occupational Pensions Authority
ICISA	International Credit Insurance and Surety Association
LoB	Line of Business
PD	Probability of Default
PIM	Partial Internal Model
PIT	Point in time
RM	Risk Margin
RR+	RecessionRisk+
SCR	Solvency Capital Requirement
TTC	Through the cycle
VaR	Value at Risk

References

1. Economic Cycle Research Institute (ECRI). What's A Business Cycle? ECRI's Website. Available online: https://www.businesscycle.com/resource-center (accessed on 30 August 2024).
2. Corporate Financial Institute (CFI). What is the Economic Cycle? CFI's Website. Available online: https://corporatefinanceinstitute.com/resources/economics/economic-cycle (accessed on 30 August 2024).
3. Claessens, S.; Ayhan Kose, M. Recession: When Bad Times Prevail. Finance & Development Magazine, International Monetary Fund (IMF). Available online: https://www.imf.org/external/pubs/ft/fandd/basics/recess.htm (accessed on 30 August 2024).
4. Claessens, S.; Ayhan Kose, M.; Terrones M. E. What happens during recessions, crunches and busts? *Econ. Policy* **2009** 24, 653–700. [CrossRef]
5. Claessens, S.; Ayhan Kose, M.; Terrones, M.E. Financial Cycles: What? How? When? IMF Working Paper WP/11/76. 2011. Available online: https://www.imf.org/external/pubs/ft/wp/2011/wp1176.pdf (accessed on 30 August 2024).
6. National Bureau of Economic Research (NBER). Business Cycle Dating. NBER's Website. Available online: https://www.nber.org/research/business-cycle-dating (accessed on 30 August 2024).
7. Bruche M.; González-Aguado, C. Recovery rates, default probabilities, and the credit cycle. *J. Bank. Financ.* **2010**, *34*, 754–764. [CrossRef]
8. Nickell, P.; Perraudin, W.; Varotto, S. Stability of rating transitions. *J. Bank. Financ.* **2000**, *24*, 203–227. [CrossRef]
9. Hamilton, D.; Su, J. High Yield's Resilience Masks Underlying Weaknessin Corporate Credit. Moody's First Quarter US Credit Review & Outlook. 2024. Available online: https://www.moodys.com/web/en/us/insights/resources/us-credit-review-and-outlook-q1-2024.pdf (accessed on 30 August 2024).
10. Directive 2009/138/EC of the European Parliament and of the Council of 25 November 2009. Available online: https://eur-lex.europa.eu/eli/dir/2009/138/oj (accessed on 30 August 2024).
11. Commission Delegated Regulation (EU) 2015/35 of 10 October 2014 Supplementing Directive 2009/138/EC of the European Parliament and of the Council on the Taking-Up and Pursuit of the Business of Insurance and Reinsurance. Available online: https://eur-lex.europa.eu/eli/reg_del/2015/35/oj (accessed on 30 August 2024).

12. Commission Delegated Regulation (EU) 2019/981 of 8 March 2019 Amending Delegated Regulation (EU) 2015/35 Supplementing Directive 2009/138/EC of the European Parliament and of the Council on the Taking-Up and Pursuit of the Business of Insurance and Reinsurance (Solvency II). Available online: https://eur-lex.europa.eu/eli/reg_del/2019/981/oj (accessed on 30 August 2024).
13. Giacomelli, J.; Passalacqua, L. Unsustainability Risk of Bid Bonds in Public Tenders. *Mathematics* **2021**, *9*, 2385. [CrossRef]
14. Giacomelli, J. The Rating Scale Paradox: An Application to the Solvency 2 Framework. *Standards* **2023**, *3*, 356–372. [CrossRef]
15. Giacomelli, J. Parametric estimation of latent default frequency in credit insurance. *J. Oper. Res. Soc.* **2022**, *74*, 330–350. [CrossRef]
16. Committee of European Insurance and Occupational Pensions Supervisors (CEIOPS). Solvency II Calibration Paper. CEIOPS-SEC-40-10. 15 April 2010. Available online: https://register.eiopa.europa.eu/CEIOPS-Archive/Documents/Advices/CEIOPS-Calibration-paper-Solvency-II.pdf (accessed on 30 August 2024).
17. International Credit Insurance & Surety Association. Comments Template on Consultation Paper on EIOPA's Second Set of Advice to the European Commission on Specific Items in the Solvency II Delegated Regulation. Feedback from ICISA to EIOPA. January 2018. Available online: https://register.eiopa.europa.eu/Publications/Comments/Oliver%20Lockwood_01_03_18.pdf (accessed on 30 August 2024).
18. EIOPA. EIOPA's Second Set of Advice to the European Commission on Specific Items in the Solvency II Delegated Regulation. EIOPA-BoS-18/075, 28 February 2018. Available online: https://www.eiopa.europa.eu/publications/eiopas-second-set-advice-european-commission-specific-items-solvency-ii-delegated-regulation_en (accessed on 30 August 2024).
19. Credit Suisse First Boston. *CreditRisk$^+$, A Credit Risk Management Framework*; Credit Suisse First Boston: London, UK, 1998.
20. Wilde, T. CreditRisk$^+$. In *Encyclopedia of Quantitative Finance*; John Wiley & Sons: Hoboken, NJ, USA, 2010.
21. Vandendorpe, A.; Ho, N.D.; Vanduffel, S.; Van Dooren, P. On the parameterization of the CreditRisk$^+$ model for estimating credit portfolio risk. *Insur. Math. Econ.* **2008**, *42*, 736–745. [CrossRef]
22. Giacomelli, J.; Passalacqua, L. Calibrating the CreditRisk$^+$ Model at Different Time Scales and in Presence of Temporal Autocorrelation. *Mathematics* **2021**, *9*, 1679. [CrossRef]
23. Giacomelli, J.; Passalacqua, L. Improved precision in calibrating CreditRisk$^+$ model for Credit Insurance applications. In *Mathematical and Statistical Methods for Actuarial Sciences and Finance—eMAF 2020*; Springer International Publishing: Berlin/Heidelberg, Germany, 2021. [CrossRef]
24. Basel Committee on Banking Supervision (BSBC). The Internal Ratings-Based Approach. Bank for International Settlements. January 2001. Available online: https://www.bis.org/publ/bcbsca05.pdf (accessed on 30 August 2024).
25. Đurović, A. Macroeconomic Approach to Point in Time Probability of Default Modeling—IFRS 9 Challenges. *J. Cent. Bank. Theory Pract.* **2019**, *8*, 209–223. [CrossRef]
26. Borio, C. The financial cycle and macroeconomics: What have we learnt? *J. Bank. Financ.* **2014**, *45*, 182–198. [CrossRef]
27. Borio, C.; Drehmann, M.; Xia D. The financial Cycle and Recession Risk. BIS Quarterly Review. December 2018; pp. 59–71. Available online: https://www.bis.org/publ/qtrpdf/r_qt1812g.pdf (accessed on 30 August 2024).
28. Borio, C.; Drehmann, M.; Xia, F.D. Forecasting recessions: The importance of the financial cycle. *J. Macroecon.* **2020**, *66*, 103258. [CrossRef]
29. Liu W.; Moench E. What predicts US recessions? *Int. J. Forecast.* **2016**, *32*, 1138–1150. [CrossRef]
30. Giesecke, K.; Longstaff, F.A.; Schaefer, S.; Strebulaev, I. Corporate bond default risk: A 150-year perspective. *J. Financ. Econ.* **2011**, *10*, 233–250. [CrossRef]
31. Wilson, T. Portfolio credit risk, I. *Risk* **1997**, *10*, 111–117. [CrossRef]
32. Wilson, T. Portfolio credit risk, II. *Risk* **1997**, *10*, 56–61.
33. Giesecke, K.; Weber, S. Cyclical correlations, credit contagion,and portfolio losses. *J. Bank. Financ.* **2004**, *28*, 3009–3036. [CrossRef]
34. Carlehed, M.; Petrov, A. A methodology for point-in-time-through-the-cycle probability of default decomposition in risk classification systems. *J. Risk Model. Valid.* **2012**, *6*, 3–25. [CrossRef]
35. Kotz, S.; Adams, J.W. Distribution of Sum of Identically Distributed Exponentially Correlated Gamma-Variables. *Ann. Math. Stat.* **1964**, *35*, 277–283. [CrossRef]
36. Florent, C.; Borgnat, P.; Tourneret, J.; Abry, P. Parameter estimation for sums of correlated gamma random variables. Application to anomaly detection in Internet Traffic. In Proceedings of the 2008 IEEE International Conference on Acoustics, Speech and Signal Processing, Las Vegas, NV, USA, 31 March–4 April 2008.
37. Feng, Y.; Wen, M.; Zhang, J.; Ji, F.; Ning, G. Sum of arbitrarily correlated Gamma random variables with unequal parameters and its application in wireless communications. In Proceedings of the IEEE 2016 International Conference on Computing, Networking and Communications (ICNC), Kauai, HI, USA, 15–18 February 2016. [CrossRef]
38. Higham, N. Computing the nearest correlation matrix—A problem from finance. *IMA J. Numer. Anal.* **2002**, *22*, 329–343. [CrossRef]
39. Italy Parliament Gives Final Approval to Austerity Plan. Reuters. September 2011. Available online: https://www.reuters.com/article/us-italy-austerity-idUSTRE78C7NF20110914/ (accessed on 30 September 2024).
40. De Grauwe, P.; Ji, P.; Steinbach, A. The EU debt crisis: Testing and revisiting conventional legal doctrine. *Int. Rev. Law Econ.* **2017**, *51*, 29–37. [CrossRef]

41. Hobelsberger, K.; Kok, C.; Mongelli, F.P. A Tale of Three Crises: Synergies Between ECB Tasks. ECB Occasional Paper Series No 305. September 2022. Available online: https://www.ecb.europa.eu/pub/pdf/scpops/ecb.op305~f9d43bd762.en.pdf (accessed on 30 September 2024).
42. Non-Performing Loans (NPLs) in Italy's Banking System. 2017. Available online: https://www.bancaditalia.it/media/approfondimenti/2017/npl/index.html (accessed on 30 September 2024).
43. The Count of Newly Non-Performing Borrowers During the Period Series Is Labelled as *TRI30529_351121441*. The Count of Performing Borrowers at the Beginning of the Period Series Is Labelled as *TRI30529_351122141*. Data Are Filtered by the ATECO 2007 Economic Sector "*F-Constructors*". Bank of Italy Statistical Database. Available online: https://infostat.bancaditalia.it/inquiry/ (accessed on 30 September 2024).

MDPI AG
Grosspeteranlage 5
4052 Basel
Switzerland
Tel.: +41 61 683 77 34

Mathematics Editorial Office
E-mail: mathematics@mdpi.com
www.mdpi.com/journal/mathematics

Academic Open
Access Publishing

mdpi.com

ISBN 978-3-7258-4246-9