economies

Special Issue Reprint

Income Distribution, Inequality and Poverty

Evidence, Explanations and Policies

Edited by
Fabio Clementi

mdpi.com/journal/economies

MDPI

Income Distribution, Inequality and Poverty: Evidence, Explanations and Policies

Income Distribution, Inequality and Poverty: Evidence, Explanations and Policies

Editor

Fabio Clementi

MDPI

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

Editor
Fabio Clementi
University of Macerata
Macerata
Italy

Editorial Office
MDPI AG
Grosspeteranlage 5
4052 Basel, Switzerland

This is a reprint of articles from the Special Issue published online in the open access journal *Economies* (ISSN 2227-7099) (available at: https://www.mdpi.com/journal/economies/special_issues/JTG6GF449G).

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

Lastname, Firstname, Firstname Lastname, and Firstname Lastname. Article Title. *Journal Name* **Year**, *Volume Number*, Page Range.

ISBN 978-3-7258-2441-0 (Hbk)
ISBN 978-3-7258-2442-7 (PDF)
doi.org/10.3390/books978-3-7258-2442-7

Contents

About the Editor

Fabio Clementi

Fabio Clementi, Ph.D., is a Professor of Economics at the University of Macerata (Italy). His main research interests focus on the size distribution of income, wealth and firms, business cycle analysis, and the empirical validation of agent-based economic models with real-world data. He has published several papers in peer-reviewed international journals and book chapters on topics related to his scientific activity and serves as a referee for various international journals. He has also been involved in a number of national and international research projects aligned with his areas of expertise. He has presented communications at many international meetings, also as an invited speaker, and has contributed as a member of the scientific and/or organizing committee of some international conferences. Most recently, he has acted as a consultant for the World Bank on matters related to income distribution and inequality.

economies

Editorial

Income Distribution, Inequality and Poverty: Evidence, Explanations and Policies

Fabio Clementi

Department of Political Science, Communication and International Relations, University of Macerata, Via Don Minzoni 22/A, 62100 Macerata, Italy; fabio.clementi@unimc.it; Tel.: +39-0733-258-2560

In recent years, there has been an increasing focus on the complex challenges posed by income inequality and poverty, both in developed and developing nations. These issues have captured the attention of economists, policymakers, and government institutions, due to their growing significance. Despite the implementation of various initiatives aimed at addressing poverty and inequality, such as the Europe 2020 Strategy and the United Nations' 2030 Agenda for Sustainable Development, progress has been limited. This can be attributed to a range of factors, including significant global events such as the Great Recession, the COVID-19 pandemic, and ongoing conflicts such as the wars in Ukraine and Gaza, which have exacerbated poverty and inequality on a global scale.

The articles presented in this Special Issue of *Economies*, entitled "Income Distribution, Inequality and Poverty: Evidence, Explanations and Policies", offer a comprehensive exploration of income inequality and poverty, shedding light on their complex drivers across various global contexts. These studies delve into a wide range of factors, from energy crises and government policies to the impacts of financial instability and global pandemics, highlighting the persistent and evolving nature of inequality and poverty in both developed and developing nations. Using diverse methodologies—including econometric models, trend analyses, and comparative cross-country studies—the contributions provide a nuanced understanding of how structural forces such as education, labor dynamics, and marketization interact to shape income distribution and social progress.

Several articles focus on specific regions, highlighting the unique challenges they face in addressing inequality. For example, Pakistan and South Africa are examined through the lenses of energy tariffs and social grants, respectively, while Central and Eastern European countries are analyzed in terms of the role of EU cohesion policies and labor market reforms. These studies collectively point to the importance of targeted policy interventions and long-term strategies to mitigate inequality, with a particular emphasis on public investment in education, social safety nets, and labor market reforms.

Other contributions take a broader, global perspective, such as the study on intergenerational educational mobility across 82 countries, or the analysis of income polarization in Europe. These papers emphasize the complex interplay between economic growth, governance, and social policies, while also recognizing the role of historical and structural factors in shaping inequality trends. Collectively, the articles in this issue offer valuable insights for policymakers, researchers, and practitioners seeking to understand and address the deep-seated inequalities that persist across regions and sectors.

The Special Issue brings together 11 original articles that delve into various dimensions of inequality and poverty across different regions and time periods. A thorough summary of the key findings and insights from each contribution is provided to guide readers in identifying the articles most relevant to their interests. Further details on each individual paper can be found in the List of Contributions.

Contribution 1 investigates how energy crises in Pakistan have deepened income inequality from 1997 to 2021. Drawing on Piketty's hypothesis, the study reveals that lower-income groups are disproportionately affected, as they spend a larger portion of their income on energy compared to wealthier segments of the population. This inequality

Citation: Clementi, Fabio. 2024. Income Distribution, Inequality and Poverty: Evidence, Explanations and Policies. *Economies* 12: 276. https://doi.org/10.3390/economies12100276

Received: 9 October 2024
Accepted: 10 October 2024
Published: 14 October 2024

is particularly pronounced in rural areas, and the paper calls for energy tariff reforms and social safety programs to mitigate these disparities.

Contribution 2 explores income inequality in South Africa between 1975 and 2017, using a Vector Error Correction Model. The study finds that increased government spending on social grants helps reduce inequality, while factors such as gross savings and population growth contribute to worsening it. Economic growth's effect on inequality is more complex, presenting mixed results. The authors recommend policies aimed at expanding social grants and fostering a culture of savings, while also addressing the challenges posed by population growth.

Contribution 3 shifts focus to the UK, analyzing when and why inequality has become a key topic in public discussions from 2004 to 2023. Using Google Trends data, the study shows that concern about inequality peaks during parliamentary elections and periods of economic instability, such as recessions. These discussions are often tied more to issues of fairness than to concerns about crime or war, particularly during periods of inflation or fluctuations in GDP.

Contribution 4 takes a global perspective, examining how educational attainment is transmitted across generations in 82 countries. The study identifies factors such as public spending on education, income inequality, and urbanization as key influences on educational mobility. Using cluster analysis, countries are categorized by their levels of educational mobility and economic development, with the author highlighting the importance of public investment in education for fostering equality of opportunity, particularly in middle- and high-income nations.

In Contribution 5, the authors investigate how the 2008 financial crisis impacted well-being across European countries over the following decade. Using Principal Component Analysis, they assess changes in multiple dimensions, such as poverty risk, material deprivation, and labor market conditions. The study reveals that different clusters of countries experienced varied impacts, with labor market issues becoming a dominant factor in the decline of well-being, especially in Southern and Eastern Europe.

Contribution 6 examines the effects of marketization, labor dynamics, and economic structures on income inequality in Central and Eastern European countries. The study highlights the roles of economic growth, education investment, labor market reforms, globalization, and governance in shaping income distribution. It suggests that targeted policies in these areas can help reduce inequality, even in the context of market-driven economies.

Contribution 7 assesses the impact of EU cohesion policies on economic growth and social progress in Central and Eastern Europe. The findings suggest that EU cohesion policies have a positive effect on both economic performance and social outcomes, particularly by reducing income inequality and poverty. However, the economic benefits take time to materialize, underscoring the importance of a long-term strategy for maximizing the effectiveness of these policies.

In Contribution 8, the authors explore the relationship between gender equality in managerial positions and wages across 22 European countries. Drawing on data from the European Structure of Earnings Survey, the study shows that while gender equality in management has mixed effects on wages—positive in the upper parts of the wage distribution, but negative in the middle and lower parts—it tends to reduce the overall gender wage gap, especially in the middle and lower segments.

Contribution 9 focuses on income polarization in 12 European countries from the early 2000s to the late 2010s. Using the relative distribution method, the study finds an increasing trend in polarization, particularly in the lower income segments. Factors such as occupational sector, education level, and area of residence are key drivers of this phenomenon. The study suggests that coordinated policy interventions could help mitigate the effects of polarization across Europe.

Contribution 10 investigates the relationship between wage inequality and operating revenues in more than 5000 Norwegian enterprises from 2008 to 2014. The study finds that wage inequality can negatively impact operating revenues, as lower wages for many

employees tend to reduce motivation and productivity. Conversely, declining revenues are shown to increase wage inequality, with the lowest earners experiencing the largest relative reductions in pay. The authors argue that a more equitable wage distribution could benefit enterprise performance.

Finally, Contribution 11 examines the impact of both the 2007–2009 financial crisis and the COVID-19 pandemic on income inequality and poverty in Greece. The study finds that both crises exacerbated inequality, with economically disadvantaged groups suffering the most. Unemployment and limited government redistribution played key roles in worsening poverty and inequality during these periods. Although pre-election periods saw temporary reductions in inequality, these effects were short-lived, and primarily benefited middle-income groups.

The contributions in this Special Issue offer a valuable foundation for understanding the key drivers of inequality and poverty across diverse regions and contexts. However, there remain several areas that invite further exploration and could significantly enrich the current discourse on inequality. One such area is the impact of climate change which, while briefly touched upon through energy crises, merits deeper investigation. Understanding how climate-related challenges disproportionately affect vulnerable communities, particularly in terms of resource scarcity and forced migration, could provide important insights.

Additionally, the role of technological advancements, such as digitalization, automation, and artificial intelligence, presents an emerging field of study. These developments are rapidly transforming labor markets, and exploring their effects on income distribution and employment opportunities, especially for those with less adaptable skills, would be a valuable contribution to the understanding of inequality.

Health inequality, another crucial determinant of economic well-being and social mobility, also deserves further research. The ways in which disparities in healthcare access intersect with income inequality, particularly during crises such as the COVID-19 pandemic, represent an area that could significantly enhance the discussions presented in this issue.

While gender wage inequality is addressed in the current collection, a deeper examination of the broader impacts of inequality on other marginalized groups, such as individuals with disabilities or ethnic minorities, would add to the understanding of the multiple barriers these groups face. Exploring how these intersecting identities influence access to education, employment, and social services could provide a more comprehensive view of inequality.

Furthermore, extending the focus from income inequality to wealth inequality would offer a more nuanced perspective on socioeconomic disparities. Wealth distribution, which includes assets such as property and stocks, is often more concentrated than income and can have lasting effects on social mobility.

Finally, the dynamics of global migration, particularly the experiences of vulnerable and undocumented migrants, offer another promising avenue for research. Understanding how migration flows impact inequality in both sending and receiving countries could yield valuable insights for policymakers and scholars alike.

In addition to exploring these thematic areas, more focused studies on the measurement of distributive phenomena, such as inequality, poverty, and polarization, are essential for advancing this field. While various approaches have been used to assess these issues, refining the methodologies and metrics for capturing the complexity and fluidity of inequality and poverty could lead to more accurate and comprehensive analyses. Improved measurement tools would also help in understanding the evolving nature of these phenomena, providing stronger empirical foundations for policy interventions.

By expanding research into these areas and refining the tools for measuring inequality, poverty, and polarization, we can develop a more holistic understanding of the global forces driving inequality and poverty. These additional lines of inquiry would complement the findings in this Special Issue, contributing to the development of more targeted and effective policy solutions.

I would like to express my sincere gratitude to all the contributing authors for their exceptional efforts in meeting the journal's standards. The success of this Special Issue in attracting a substantial number of contributions reaffirms the potential of open access publishing in our field. I am sincerely thankful to the numerous reviewers whose invaluable feedback has greatly enhanced the quality of the published papers. Finally, I would like to extend my deep appreciation to the team of assistant editors, especially Ms. Adore Zhou, for their outstanding support.

Funding: This research received no external funding.

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

List of Contributions

1. Hussain, J.; Siyal, S.; Ahmad, R.; Abbas, Q.; Yitian, Y.; Jin, L. Impact of Energy Crises on Income Inequality: An Application of Piketty's Hypothesis to Pakistan. *Economies* **2024**, *12*, 259. https://doi.org/10.3390/economies12100259.
2. Seabela, M.; Ogujiuba, K.; Eggink, M. Determinants of Income Inequality in South Africa: A Vector Error Correction Model Approach. *Economies* **2024**, *12*, 169. https://doi.org/10.3390/economies12070169.
3. Seip, K.L.; Sandnes, F.E. The Timing and Strength of Inequality Concerns in the UK Public Debate: Google Trends, Elections and the Macroeconomy. *Economies* **2024**, *12*, 135. https://doi.org/10.3390/economies12060135.
4. Valentini, E. Patterns of Intergenerational Educational (Im)Mobility. *Economies* **2024**, *12*, 126. https://doi.org/10.3390/economies12060126.
5. Croci Angelini, E.; Farina, F.; Sorana, S. The Impact of the Great Recession on Well-Being across Europe Ten Years On: A Cluster Analysis. *Economies* **2024**, *12*, 115. https://doi.org/10.3390/economies12050115.
6. Davidescu, A.A.; Lobonţ, O.-R.; Nae, T.M. The Fabric of Transition: Unraveling the Weave of Labor Dynamics, Economic Structures, and Innovation on Income Disparities in Central and Eastern Europe Nations. *Economies* **2024**, *12*, 68. https://doi.org/10.3390/economies12030068.
7. Davidescu, A.A.; Nae, T.M.; Florescu, M.-S. From Policy to Impact: Advancing Economic Development and Tackling Social Inequities in Central and Eastern Europe. *Economies* **2024**, *12*, 28. https://doi.org/10.3390/economies12020028.
8. de Castro Romero, L.; Barroso, V.M.; Santero-Sánchez, R. Does Gender Equality in Managerial Positions Improve the Gender Wage Gap? Comparative Evidence from Europe. *Economies* **2023**, *11*, 301. https://doi.org/10.3390/economies11120301.
9. Fabiani, M. Unraveling the Roots of Income Polarization in Europe: A Divided Continent. *Economies* **2023**, *11*, 217. https://doi.org/10.3390/economies11080217.
10. Aarstad, J.; Kvitastein, O.A.; Wage Inequality's Decreasing Effect on Enterprise Operating Revenues. *Economies* **2023**, *11*, 178. https://doi.org/10.3390/economies11070178.
11. Petrakos, G.; Rontos, K.; Vavoura, C.; Vavouras, I. The Impact of Recent Economic Crises on Income Inequality and the Risk of Poverty in Greece. *Economies* **2023**, *11*, 166. https://doi.org/10.3390/economies11060166.

economies

MDPI

Article

Impact of Energy Crises on Income Inequality: An Application of Piketty's Hypothesis to Pakistan

Jibran Hussain [1], Saeed Siyal [2,*], Riaz Ahmad [3], Qaiser Abbas [4], Yu Yitian [5] and Liu Jin [5,*]

[1] Riphah Institute of Public Policy, Riphah International University, Islamabad 46000, Pakistan; jibranhr@gmail.com
[2] Human Capital Research Center, College of Business and Economics, United Arab Emirates University, Al-Ain P.O. Box 15551, United Arab Emirates
[3] Bahria Business School, Bahria University, Islamabad Campus, Islamabad 54000, Pakistan; riaz.buic@bahria.edu.pk
[4] Bahria Business School, Bahria University, Karachi Campus, Karachi 73500, Pakistan; qaiserabbas51412@gmail.com
[5] School of Humanities and Social Sciences, Beijing Institute of Technology, Beijing 100081, China; yuyitian93@163.com
* Correspondence: saeedsiyal@gmail.com (S.S.); liujinedu@bit.edu.cn (L.J.)

Abstract: In Pakistan, the majority of people have access to energy supplies. However, people who are underprivileged, below the extreme poverty line, or part of the middle class often spend disproportionate portions of their income on energy supplies and services, to some extent because of higher upfront prices for energy supplies, expensive products, and expensive imported appliances. The nonavailability of low-cost energy supplies is mainly affecting underdeveloped regions that have mostly low-income households. We used the dynamic ordinary least squares method to look at the impact of the energy crisis on income inequality from 1997 to 2021. The results show that the energy crisis exacerbates income inequality as low-income groups end up spending more significant shares of their income on energy products, supplies, and services than higher-income groups. Fair and equal access to energy supplies and services is less likely to reduce income inequality if prices are not cost-efficient. Cautious deliberation regarding the structure of energy tariffs is inevitable; at the same time, safety nets and social security programs for the poorest groups need to be expanded. At this stage, the aim is to target energy prices that will achieve the objectives of reducing polarity and increasing real income.

Keywords: energy crisis; income inequality; Piketty's hypothesis

Citation: Hussain, Jibran, Saeed Siyal, Riaz Ahmad, Qaiser Abbas, Yu Yitian, and Liu Jin. 2024. Impact of Energy Crises on Income Inequality: An Application of Piketty's Hypothesis to Pakistan. *Economies* 12: 259. https://doi.org/10.3390/economies12100259

Academic Editor: Angeliki N. Menegaki

Received: 16 July 2024
Revised: 16 August 2024
Accepted: 5 September 2024
Published: 24 September 2024

1. Introduction

The distribution of resources has always been one of the central points of socioeconomic and sociopolitical policy-making history, as well as the history of class struggle, a struggle for resource acquisition and exploitation (Marx and Engels 2008). This article delves into this critical yet overlooked area in the context of Pakistan, aiming to contribute to the existing literature by examining the historical and contemporary aspects of resource distribution. Here, we define the term 'energy crisis', central to our discussion, as the significant lack of energy supplies and accessibility, particularly affecting the lower and middle classes and exacerbating income inequality. The impact of the Industrial Policy (1948) and the Green Revolution (1959), though responsible for modest economic growth from 1948 to 1966, laid the foundations for the vast disparity in society with the emergence of industrial capitalist, feudal, and agricultural capitalist classes. The period from 1979 to 1989 marked an adjustment toward free enterprise development and transformation; as a result, the capitalists flourished. The decade of 1990–2000 in the history of the country and the accompanying political volatility did not add much to the redistribution of resources

or to lifting the lower and middle classes. This ongoing issue of resource distribution and its historical roots needs in-depth investigation. Moreover, the political stability crisis was vital in contributing to income inequality, as well as the energy crisis, during this period. The political climate was not stabilized through any remarkable or notable institutional capacity-building programs in the years 2000 to 2007; rather, ad hoc adjustment policies continued to be enforced as a popular method of running state affairs. The energy fabric and system collapsed during the recession, led by the energy crisis in 2007, and there was not much that political and economic institutions could offer to society (Zaidi 2015).

The country is still struggling with the aftermath of the policy-incapacitated outcomes of the past. As a result, inequality kept gradually increasing; the share of total income of the poorest 20 percent decreased during the period from 1997 to 2021, while the share of the wealthiest 20 percent increased. The share of the poorest 20 percent increased slightly during the 1980s. However, their share started to decline again in the 1990s and reached 6.37 percent by 2004. On the other hand, the share of the wealthiest 20 percent increased during the period from 1997 to 2021 and reached 50.02 percent in the period from 2009 to 2010 (World Bank Group 2014). Although the average income level in Pakistan is slightly increasing, along with economic growth of almost 5 percent, the gap between socioeconomic classes has not improved significantly. In this context, this article examines the factors beyond income levels that shape equality in society, with a special focus on the availability and affordability of energy resources. This article shows that there are factors other than income levels and economic growth that have significant roles in shaping equality in society. Two of these critical factors are affordable energy services and access to energy resources. The growing and creeping crisis in the energy sector over the years, as a result of persistent political uncertainty in the country, has been decreasing the capacity of the masses to access energy resources; for Murtaza and Faridi (2015), it has increased energy poverty, which reflects the persistency of the lack of energy supplies and the inability of the masses to access resources. At the same time, the limited availability of energy supplies and services has been reducing the real income of the masses over the years.

According to Cheema and Sial (2012), this increasing income inequality despite increasing economic growth rates is due to the unfair distribution of resources among provinces and the biased distribution between rural and urban areas. In the case of rural–urban polarization, a similar conclusion was reached by Awan et al. (2013), in that the ability to access energy services in rural areas is relatively lower than in urban areas. Moreover, Haq and Shirazi (1998) concluded that, in terms of nonfood expenditure, economic welfare levels in rural areas are comparable to those of the urban sector, and they proposed that policies should be directed toward enhancing the expenditure capacity of the poorest groups. Nwosu et al. (2018) showed that nonfood expenditure is a major source of inequality in household consumption expenditures in both urban and rural areas and that variables such as living in rural areas, household size, type of household dwelling, and household dwelling characteristics account for significant proportions of inequality in food and nonfood expenditures. Murtaza and Faridi (2015) blamed income polarization for expanding income poverty and energy poverty, which reflects the decreasing consumption expenditures.

Pakistan is among the weakest countries in the world in terms of the Sustainable Energy Development Index (SED*I*), the Human Development Index (HD*I*), and the Energy Poverty Index (EPI) tested globally. This implies that the average household in Pakistan has to spend a large portion of its income to acquire energy supplies and services every time there is an energy crisis or oil prices increase. As over-dependency on external energy supplies shifts wealth from the local economy to the foreign, leaving less for the locals, the growth in household final consumption expenditure has decreased over the years in the country (Iddrisu and Bhattacharyya 2015). Kuznets (1955) contributed to the issue of income and resource distribution by empirically analyzing income inequality. He believed that countries face serious inequality issues when the economy is in the 'take-off' stage

and moving toward industrialization; however, the issue of inequality stabilizes after the achievement of the steady-state level.

Piketty (2014), on the other hand, dived into the wealth and tax data for France, Germany, Japan, the United Kingdom, and the USA, from 1810 to 2010, to study the resource distributions and income inequalities in these countries. He believes that the concept of inequality in the field of economics is insufficient, especially while using the relationship between the economic growth rate and income shares. He expands his detailed discussion and proposes that, if the rate of return from capital (r) exceeds the growth rate of the economy (g) over a steady period, it gives rise to the problem of income inequality (Piketty 2015). Despite expanding the critique on his theory of income inequality, he proposes that, in addition to the $r > g$ hypothesis for explaining income inequality, institutional changes and political shocks remain as integral and indigenous parts of this income inequality function (Piketty 2015). The extension of Piketty's hypothesis presents a valid fit in the case of Pakistan, where political instability over the years has been responsible for this reduction in institutional capacities; as a result, despite increasing growth rates and income levels, income inequality is on the rise, especially after the 1990s. Lakner (2016) noted that the crucial aspects of measure Piketty's (2014) that make it more relevant and appropriate are as follows: First, it does not underestimate the incomes and expenditures of wealthy households, as Atkinson et al. (2011) feared that most of the household surveys do a poor job capturing the incomes and consumptions of the richest. Second, it captures the capital incomes and profits of the entrepreneurial process, which household surveys in developing countries fail to incorporate (Alvaredo and Gasparini 2015). Third, consumption surveys, in use primarily in underdeveloped and developing countries, tend to undermine the actual picture of the living standards of the wealthy class, who tend to save more than the bottom fractions of a society (Aguiar and Bils 2015). Only 40 percent of the total population of Pakistan has access to clean fuel and technologies for cooking, and this is biased toward urban areas, where almost 98 percent of consumers have access to electricity, whereas only 90 percent of consumers in rural areas have the same access (World Bank 2018). The access, availability, and affordability of energy supplies, services, and technologies have wide disparities between the urban and rural, rich and poor, and haves and have-nots in Pakistan. This increasing inequality in energy resource distribution, energy supply disparities (Mahmood and Shah 2017), and income and energy poverty (Murtaza and Faridi 2015) in the presence of a persistent, unstable political system is causing increasing income inequality in Pakistan (Shehzadi et al. 2019).

The existing literature provides a foundation for understanding the relationship between energy access and income inequality. Studies have shown that unequal access to energy can exacerbate income disparities (Bazilian and Yumkella 2015; Bouzarovski and Herrero 2017). In Pakistan, energy poverty has been identified as a significant issue affecting low-income households (Mahmood and Shah 2017). This study builds on these findings by applying Piketty's hypothesis of resource distribution to the context of Pakistan's energy crisis.

The Kuznets curve has been a popular benchmark for understanding the relationship between economic growth and income inequality. This popularity was particularly notable during the mid-20th century.

Acemoglu and Robinson (2013) provide a deeper insight into situations in which there is an abundance of energy but high inequality. Their work suggests that inclusive political institutions are necessary to ensure equitable resource distribution.

Atkinson's contributions, particularly his previous research, have similar conclusions about the relationship between inequality and resource distribution. This highlights the need for equitable energy policies.

Piketty's explanation of the correlations between resource distribution, growth, and income inequality suggests that unequal resource distribution leads to increased inequality. Resource distribution, in this context, includes access to energy, which is essential for economic growth and reducing inequality. Grunewald et al. (2014) imply that a reduction

in CO_2 emissions, energy usage, or energy supply increases the difference between energy supply and demand, causing energy crises. Shahbaz (2010) tested Kuznets' theory in the context of energy and income inequality in Pakistan. The measure of income inequality used in these studies differs from Piketty's r-g model. (See Figure 1)

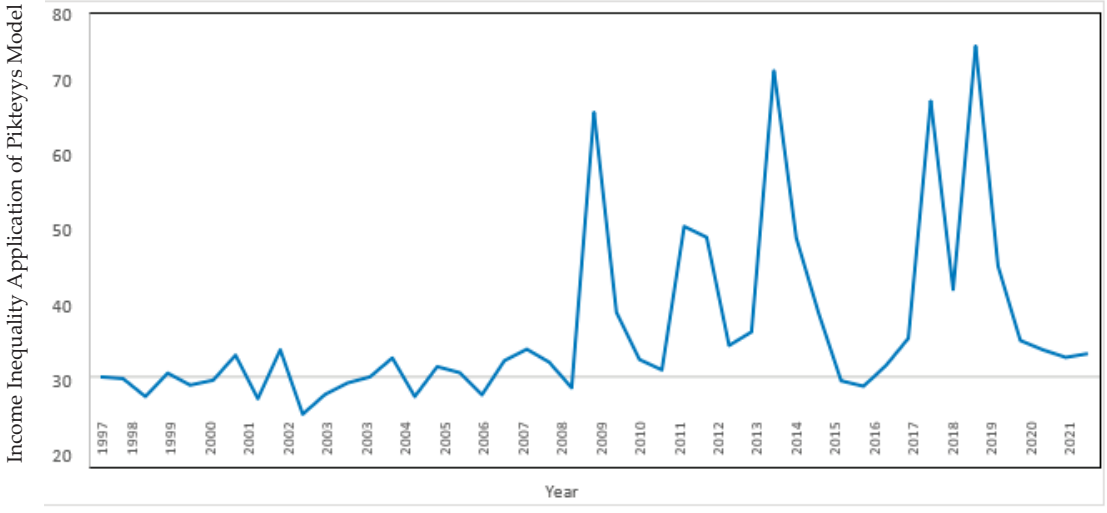


Figure 1. Application of Piketty's hypothesis on income inequality, 1997–2021.

A comparison of rural and urban areas shows that 93 percent of the urban population compared to 63 percent of the rural population in Pakistan have access to electricity (Mahmood and Shah 2017). The misery of the energy disparity and deprivation reached an alarming situation, with 19 percent of the urban and 71 percent of the rural populations not being able to afford energy services or supplies. Despite increasing average income levels in Pakistan, along with an average economic growth rate of almost five percent, household consumption expenditure is decreasing and income inequality is barely improving (Awan et al. 2013). It is evident, from the fact that, despite an increasing per capita income, the Gini coefficient increased from 30 in 1997 to 37 in 2021, and household consumption expenditure is showing a decreasing trend (World Bank 2022). The impact of the energy crisis on the rural poor is significantly more severe than that on the urban masses, and the underprivileged and the poor are spending a more significant share of their incomes on energy supplies. The energy crisis is contributing to aggravated income inequality by widening the gap between the haves and the have-nots in the country, and hampering the ability of masses to have access to energy services and supplies, by reducing the real income of the lower classes, and thus their capacity to live.

This paper makes a significant contribution by linking historical and contemporary issues of resource distribution in Pakistan with income inequality, particularly through the lens of the energy crisis. By examining the impact of policies like the Industrial Policy (1948) and the Green Revolution (1959), and their role in fostering socioeconomic disparities, the paper highlights how historical decisions have perpetuated income inequality and energy poverty. It builds on the existing literature by applying Piketty's hypothesis on resource distribution and income inequality to Pakistan's context, demonstrating that despite economic growth, the unequal distribution of energy resources exacerbates socioeconomic divides. This study underscores that beyond income levels and economic growth, access to affordable energy services is crucial in shaping societal equality. By integrating various theoretical perspectives, including Kuznets' curve and Piketty's r-g model, with empirical data on energy access and income inequality, the paper offers a nuanced analysis of how

persistent energy crises and political instability contribute to worsening income inequality in Pakistan.

The rest of this paper is as follows: Section 2 contains the literature review; Section 3 covers methodology; and Section 4 covers data analysis followed by conclusions and recommendations for future directions.

2. Literature Review

Income inequality is one of the most debated topics in the literature on socioeconomic welfare. The theory of income inequality reflects the fundamental issues of resource distribution and share, and the access of different fractions of society to national and collective resources. In every era, prominent researchers have presented their theories on income inequality, such as Smith (1776) who provided an essential and significant foundational understanding of income inequality through the concept of resource distribution between the laborers, the capitalists, and the landowners. He theorized that the landowner becomes richer with the growth of free markets and society as an consequence of the laborers hard work, this is the foundation of inequality in terms of resource distribution and the exploitation of the laboring class. The foundational work of Smith paved the way for Marx (1818–1883) to propose his famous theory of socioeconomic inequalities. He stated that the capitalist exploitation of resources is responsible for polarizing society and motivates the masses toward revolution as a form of class struggle. His thesis promised equality and justice in society and cured the issue of inequality with the prevalence of a dominant socialist society. On the contrary, Skocpol (1977) believed in the fundamental forces of market stabilization as a solution to the polarities in resource distribution and prevailing income inequalities. Wherein, Polinsky (1974) also conceded the possibility of fair distribution of resources through a theory of equity in factors of production. On the other hand, Keynes (1973) portrayed a pessimistic view of the consequences of income inequality and cautioned about the more significant damage that inequality could bring to society and the economy. Energy supply is a crucial factor for social development and growth, as well as advancements and transformations toward modernization. Limiting the supply of energy would create a crisis in societies. The sustainable flow of energy and energy use is fundamental for industrial economies and social organizations. Therefore, energy deprivation is a 'threat multiplier' and 'threat instigator' in military parlance; yet, unfortunately, it took several years for the international community and organizations to acknowledge the importance of energy sustainability and security. This energy deprivation is considered to be a clear catalyst for unrest, instability, youth unemployment, and urbanization in societies.

Income inequality is an essential determinant of socioeconomic welfare, a vital theoretical affair for policymakers and economists. It reaches resource economics with the hope of determining the redistribution of scarce resources in a way that could reduce inequalities in society. The theory of inequality by Kuznet (1955) remains a popular benchmark to measure the level of welfare through income distribution among different income fractions in society. Notwithstanding the popularity of Kuznets's thesis, the application of his measure of income inequality to developing countries raises valid and contradicting questions; most such objections are debatable, yet much is left unexplained by his theory, such as cases in which urbanization, industrialization, increasing improvements in energy resources and economic growth occur side by side along with increasing income inequality, such as in Russia (Lyubimov 2017). Similarly, Acemoglu and Robinson (2013) explained the resource richness (energy) of countries (such as Nigeria and Middle Eastern countries) experiencing increasing inequalities, thus showing that resource distribution is not justified or rational. The problem of income inequality was addressed in the groundbreaking work of Atkinson et al. (2011). He proposed a solution to the measurement of income inequality and the distribution of income, consumption, and wealth in terms of subjective social function. He measured income inequality within income distribution through 'social loss', which can be reflected through income, consumption, or wealth. Therefore, the debate

generated about socioeconomic welfare discussed earlier in this section is justified herein to explain the concept and measurement of income inequality. He believed that, as a traditional measure of income inequality, the Gini coefficient fails to explain the welfare aspects imbibed within income inequality, such as the social loss in unjust resource distribution and income inequalities. Therefore, the Gini coefficient cannot be trusted as an ultimate, complete, and reliable measure of income inequality. The concept of social loss first originated in the work of Dalton (1920), who defined income inequality as a loss of socioeconomic welfare as a result of the unjust and uneven distribution of resources and incomes within a society.

Sen (1992) re-examined the concept and measurement of income inequality on the paradigm of capability functions such as elementary function (health, nourishment, and shelter) and social function (self-respect and community life). The realization of these sets of capabilities, as preferred by any individual, reflects his/her capability and standing in society. Therefore, the failure to provide access to and the availability of these functions for a different fraction of society reflects the prevailing level of income inequality, which hinders human freedom to enjoy the basic needs and wants. Hence, his theory explained inequality in terms of social welfare, where people compromise their ability to retain a respectable justified level of life due to nonavailability, a lack of access to collective resources, and a lack of capability. This theory enhances the concepts of social loss and socioeconomic welfare while expanding the definition and scope of the concept of income inequality. However, the promising contemporary theory of inequality from Piketty (2014, 2015) proposed a whole new dimension of income inequality and filled gaps left by earlier theories, especially in the context of macroeconomic dynamics. The proposed theoretical basis is that, when the rate of return on capital (r) is higher than the economic growth rate (g) over a reasonable period, it leads toward inequality in the society ($r > g = IE$). The theoretical architecture of this theory is based on the primary and fundamental laws of capitalism. For example, the first law says that the product of the capital–income ratio (β) and return on capital (r) are equal to the share of capital in the national income (α). The second law of capitalism asserts that β is the outcome of the ratio between national saving (s) and economic growth (g). Hence, the rate of return on capital is equal to the product share of capital within the national income and ratio between economic growth and the national savings ($r = \alpha * g/s$).

Piketty's successful explanation of the correlations between resource distribution, growth, and income inequality over a greater extent of time gives a new dimension to the field. The major concern of the theory is that rising inequalities lead the wealthy class to influence the political and economic institutions in their favor; this aspect of the theory has already gained theoretical endorsement by Acemoglu and Robinson (2013) in that extractive economic institutions favor the elite to obtain personal benefits from the state and weaken the institutions' and nations' failures despite having abundant resources. Galvin (2019) proposed that countries with weak institutions cannot prevent the impact of external shocks, such as energy shocks, and such economies are prone to experiencing an economic collapse. Hence, unjust resource distribution (energy) leads to income inequality. The theoretical base between energy used/energy supply and income inequality is not widely explored or researched. However, some researchers have used carbon dioxide (CO_2) emissions as a proxy for energy supply and developed a theoretical relationship with income inequality to lay down foundations for their theories, such as the 'equality hypothesis', which states that income inequality could be positively associated with CO_2 emissions, implying that energy-use enhancement hints toward the inequality of energy supplies and, therefore, income inequality (Boyce 1994). On the contrary, some theorists believed that increasing CO_2 emissions play a fair role in bringing inequality downward, and are thus negatively associated with income inequality (Heerink et al. 2001).

Similarly, the comprehensive theoretical justification for the relationship between CO_2 emissions/energy usage and inequality proposed by Grunewald et al. (2014), that, in the case of developing and underdeveloped countries, CO_2 emissions are negatively associated with income inequality, thus implies that more energy being used in developing countries

would decrease income inequality, as the amount of energy used is an indicator of economic growth and development. Hence, an increase in energy supplies improves economic growth, which therein extends to decreasing inequalities. The authors included Pakistan in their analysis, which, therefore, implies that a reduction in CO_2/energy usage/energy supply increases the difference between energy supply and demand, which causes energy crises. This crisis puts downward pressure on economic growth, and the decreasing economic growth deteriorates the resource distribution and income inequality in Pakistan. This theory, therefore, provided the foundations for exploring the role of an energy crisis in explaining the variations in income inequality in the case of Pakistan. On the other hand, Khan and Heinecker (2018) laid the foundations for a theoretical relationship between income inequality and energy supply, while reinvigorating the debate on energy efficiency and income disparity. Their proposed theory states that an increase in energy use is associated with increasing disparity and is the result of an increase in the energy cost of having 'good regulation' of the system. As the system becomes unequal, the supply of energy resources and services becomes increasingly biased toward the high-income fractions of society, where the top 20 percent of the wealthiest fraction have more access to and use of energy than the rest. The links between energy and reducing inequalities may be seen most clearly in the context of access to energy, but there are also cases of energy poverty related to fuel poverty. In situations in which people have access to energy, it is often the poorest that end up using disproportionate shares of their income to pay for energy, in part because the higher upfront costs of investments in energy services are more difficult to bear for low-income households (Simcock et al. 2017). This thus validates the claim of Piketty (2015) that when resource distribution favors the capitalists more than the majority, the return on capital increases more than growth, hence, creating income inequality.

Energy supply is a crucial factor in social development and growth and the advancement of the transformation of society toward modernization. A limit on the supplies of energy would create a crisis in societies. The sustainable flow of energy and energy use is fundamental for industrial economies and social organizations. Therefore, to avoid the energy crisis, industrial societies are requisite to find alternative means of energy, York et al. (2004). Access to energy resources and supplies is a prerequisite for development, such that the distribution of energy supplies and its access may have social repercussions and cause economic inequalities. Awan et al. (2013) argued that the majority of households rely on fuel consumption and noted that they face almost a 55 percent deprivation in terms of energy access, especially rural households. While reporting the bearing of the energy crisis by society, Arslan et al. (2014) noted that, with just the nonavailability of CNG fuel, the social and economic wellbeing of people in Pakistan is affected adversely. Moreover, they showed that the energy crisis is responsible for affecting the lifestyles of people in general. Similarly, Simcock et al. (2017) reported that the energy crisis has adversely affected all spheres of the wellbeing of people and has escalated the cost of living, especially for the low-income groups.

Equality and welfare both rely on a sustained and stable energy supply to all. Economic growth and development is a step toward eradication of energy deprivation and inequality, and without sustainable energy supplies, it is difficult to reach the targets of economic development. The gap between the haves and have-nots in terms of access to modern energy services is widening day by day (World Bank 2013). Economic development has proven to be a significant necessary condition to enhance the energy use of the masses; thus, it is a necessary condition to reduce an energy crisis experienced by any society, York et al. (2004). Moreover, energy is the most crucial and integral part of production processes and services that enable societies to move forward on the path of growth, toward development and prosperity. Access to energy resources is a serious global issue. Today more than half of the world's population is deprived of access to energy, energy-related services, appliances, and facilities. This deprivation is a significant hurdle and creates restraints for creating jobs, business opportunities, access to health, and education (World Bank 2013). However, the situations and scenarios in developing and underdeveloped coun-

tries such as Pakistan are getting worse. Moreover, countries with humongous populations and population growth rates face energy shortages. Some of the features of developing and overpopulated countries are increasing population and urbanization. Alongside other benefits, access to energy services and supplies is better, which attracts the rural population to migrate. In his cross-country study, Liddle (2013) found that urbanization increases with energy use.

Today more than one billion people around the world have no access to electricity. This situation is forecasted to be worse by 2030, when this number will increase by as much as two-fold. The conditions in underdeveloped and developing countries are particularly terrible (Toman and Jemelkova 2003). Moreover, rural electrification has been proven to reduce energy poverty, which consequently improves energy equity and, thus, reduces inequality. Furthermore, the amount required to mitigate the gap between the haves and have-nots in terms of access to energy services, as well as to provide the deprived masses around the world with necessary energy supplies and services, is almost USD 40 billion annually, which will rise continually until 2030 worldwide (Schroeter 2013). Pao et al. (2014) stressed that global initiatives are inevitable for the provision of necessary energy supplies and services to produce opportunities for growth and prosperity for underprivileged and deprived people. Unfortunately, the masses with the least access to resources cannot lift themselves out of the abyss of misery, poverty, and income inequality without access to modern energy supplies and services. The reliance on expensive oil and fuel is not going to cater to this severe issue affecting the poor of the world. Even though the solution to this crisis of income inequality and uneven access to energy resources has been proven to lead to massive progress and development in renewable energy technologies and services, the adaptation of such technologies to developing and underdeveloped nations is near nonexistent, which is a sustained threat of increasing income inequality.

The bulk of energy sources being used around the globe are not sustainable (International Atomic Energy Agency 2005); therefore, energy deprivation is a 'threat multiplier' and 'threat instigator' in military parlance; yet unfortunately, it took several years for international communities and organizations to acknowledge the importance of energy sustainability and security. This energy deprivation is considered a clear catalyst for unrest, instability, youth unemployment, and urbanization in societies (Bazilian and Yumkella 2015). The literature on energy usage/energy supply and income inequality is scarce, primarily due to the ambiguous theoretical paradigm between the two variables (Berthe and Elie 2015), and empirical studies on the energy crisis and income inequality are almost nonexistent, around the world in general, and in Pakistan in particular. Therefore, CO_2 emissions can be considered as a proxy of energy use, owing to the fact that more than 77 percent of CO_2 emissions are from energy sources (EIA, USA 2019; Ceyhan and Saribas 2022). As such, Shahbaz (2010), while exploring the environmental Kuznets curve and the role of energy consumption in Pakistan, stated that energy consumption increases CO_2 emissions. As per Grunewald et al. (2014), increasing CO_2 reduces inequality in developing countries. Therefore, with increasing CO_2 emissions, income inequality may reduce in Pakistan.

In purview of this, an increase in energy consumption signals an increase in energy supplies, and if access to energy services have increased, and if this expansion includes the rural and deprived areas, it certainly provides room for us to infer that an energy supply has become available to the deprived. Therefore, in order to find to a shred of justified empirical evidence and to investigate a plausible association between energy use and income inequality, in many studies CO_2 emissions can be relied upon and inferred to be the energy usage. As such, Heerink et al. (2001) tested the relationship between CO_2 emissions and income inequality in 65 countries and found that a negative association between income inequality and CO_2 emissions exists. These results insinuated that the increase in CO_2 emissions reduces income inequality. Similar results are also endorsed by Ravallion et al. (2000) in their study using data for 42 countries from 1975 to 1992, and they proved that there is a static trade-off between CO_2 emissions and income inequality. On the

contrary, a rather comprehensive and more convincing empirical study in this aspect was completed by Grunewald et al. (2014), from 1980 to 2008, using a sample of 90 countries to test the relationship between CO_2 emissions and income inequality. They concluded that countries with low- and middle-income statuses possess a negative relationship between CO_2 emission and income inequality, and, on the contrary, this relationship is positive in the case of developed and high-income countries. Reaffirming the above results, in their study, Grunewald et al. (2017) found that, for low- and middle-income economies, higher income inequality is associated with lower carbon emissions, while in upper-middle-income and high-income economies, higher income inequality increases with per capita emissions. Since Pakistan is included in the group of low-income countries by the authors, it is implied that there is a negative relationship between CO_2 emissions and income inequality.

The studies mentioned above and their results imply that income inequality decreases as CO_2 emissions increase. And the increase in CO_2 emissions is the imposition of the fact that more energy is being used at large (EIA, USA 2019). Therefore, the possibility that more people are using more energy and have greater access to energy services with the enhanced capacity to buy energy resources may indicate a reduction in income inequality. Hence, it provides an avenue of justification that energy use is a vital tool for reducing income inequality. Moreover, the reduction in CO_2 emissions signals in a drop in energy use. A sustainable supply of energy resources and services is a necessary condition for the development stages of societies. However, with the increasing utilization of energy resources, and a growing population, the stock and resources of energy supplies worldwide are decreasing significantly, raising fears for greater energy deprivation and polarity among nations and risking the share of energy for future generations (Sahir and Qureshi 2007). As such, access to energy supplies and services affect the freedom of millions around the globe to access economic opportunities. One of the reasons for areas of the world having remained underdeveloped is the lack of energy supply and services to initiate economic activities (Bazilian and Yumkella 2015). Income inequality, energy poverty, and poverty move in the same direction, as changes in energy consumption's impact are a reliable indicator of prosperity. This is why the negative correlation between modern energy services and energy deprivation is well established and proven. In order to reduce income inequality and energy deprivation, access to modern energy services by the masses must be improved (International Energy Agency 2017).

Awan et al. (2013) investigated the situation of energy deprivation and energy poverty in Pakistan by using the Multidimensional Energy Poverty Index (MEPI). They found a high intensity of energy deprivation and poverty throughout the country. The comparative analysis and results between rural and urban populations showed that the urban population in the country enjoys greater access to energy services than the rural population. As much as 71 percent of the population in rural areas of Pakistan are deprived of energy services and supplies compared to 29 percent in urban areas. This situation is a reflection of the prevalence of disparities in this society, in which urban populations not only have greater access to energy supplies but also socioeconomic opportunities to explore and with a higher per capita income compared to the rural populace. This also suggests that the rural populace comparatively experience a higher-pressure energy crisis, first based on their limited access to energy supplies and services, and, secondly, their ability to buy more energy when it is expensive remains low due to their income. Therefore, the persistent energy crises over the years has contributed to widening these gaps in polarity and income inequality in the country.

This study employs the dynamic ordinary least squares (DOLS) method to analyze the impact of the energy crisis on income inequality from 1997 to 2021. The DOLS method is chosen for its ability to handle endogeneity and serial correlation issues, providing unbiased and efficient estimates (Lütkepohl 2001). This method is particularly relevant for our study as it allows for the incorporation of long-term equilibrium relationships among variables, which is essential for understanding the persistent effects of energy crises on income inequality.

Although, MEPI as an index for energy deprivation and poverty, it is also an excellent way to measure the income inequality derived through energy supply and services; however, the authors believe that it does not integrate all of the elements of sustainable energy. As such, advancement in the field of energy economics is required. Mahmood and Shah (2017) also conceived similar results, while extending the application of MEPI to the rural and urban areas of Pakistan. Their study remained focused on a comparative analysis and investigation of the differences in energy deprivation and access to energy resources and services between rural and urban areas of the country. They found that, on average, a household in Pakistan is 26.4 percent deprived of essential energy services and supplies. The alarming nature of this situation is that this deprivation is chronic and persistent, with the deprivation in rural areas being massively greater compared to the level of energy deprivation in urban areas of the country. With these results, it can be argued that the polarity and deprivation in access to primary energy supply and services are the foundational cause of persistent income inequality and the energy crisis in the country. They believe that the lack of governing wisdom and political will are responsible for bringing society to this edge. Murtaza and Faridi (2015) already produced similar results and showed that the progress in energy development in Pakistan increased during the first half of this century but at a plodding pace. Unfortunately, even this slow and meagre energy development and growth started decreasing after the energy shocks and crises of 2007 and 2011, resulting in increased energy poverty, income inequality, and disparities of the country. The level of income inequality is rooted in the overall economic progress and development of a country. An increasing economic performance is considered to have an impact on the level of earnings of households, which eventually transforms into households' increasing ability to access to modern services and energy resources. Such an idea is proposed by Kuznet (1955), who believed that structural adjustment of economic growth can reduce the deprivation and income inequalities in society, as society crosses different levels of growth and development. He further contemplated that the very early and initial stages of economic growth yield income inequality, and later, when a steady state is achieved, income inequality is potentially reduced, based on the premise of industrialization, urbanization, and the trickle-down effect of economic growth to the lowest fraction of the society. Over the years, economists all over the world have studied the application of Kuznets theory in different scenarios without considering the collected historical data or improving the empirical validity of the theory to their respective societies (Piketty 2014). Shahbaz (2010) witnessed a similar application of Kuznets's theory of income inequality in Pakistan and showed that economic growth leads to a reduction in inequality. However, the author failed to align his results and outcomes with the growth stages of the underpinning theory. On the contrary, Zouhaier and Karim (2012) found a negative effect of economic growth on inequality and proposed the need for a better methodological understanding to obtain more accurate results. Sharafat et al. (2014) reached similar results and found that increasing economic growth is associated with the increase in income inequality in Pakistan.

Sen (1992) raised a point on the moral validity of researchers' approaches toward studying inequality rather than using objective-driven research. Similar yet rigorous results are forwarded by Checchi and García-Peñalosa (2008). They studied 21 countries and reached the conclusion that labor market institutions are the prominent driving force in the determination of inequality; a more reliable institution would result in decreasing inequality.

Sustained and sufficient energy resource distribution is predominantly agreed to rely on the factor of socioeconomic prosperity. However, energy resource allocation, efficiency, and conservation have long been critical elements in the energy policy dialogue. They have taken on renewed importance as concerns about global climate change and energy security have intensified in a country where 51 million or more people are still without electricity, with the national electrification rate being almost 73 percent (World Bank Group 2014). Many advocates and policymakers maintain that reducing the demand for energy is essential to meeting the challenges of energy deprivation and inequalities, and analyses tend to find that reductions in demand can be cost-effective means of addressing the

concerns of a shortage in energy supplies in the country and hence can curtail the ongoing energy crisis, being known as energy conservation plans. However, there is a difference in reducing demand and curtailing it. For instance, rural areas in Pakistan barely have a 63 percent electrification rate compared to 90 percent in urban areas, showing the inequality in energy resource distribution. Therefore, any conservation program that reduces energy demand indifferently between the urban and rural masses would hamper the social and economic activities of the country and eventually increase the threat to socioeconomic growth and welfare (World Bank Group 2014). The energy resource distribution and income inequality hypothesis proved that, in the case of developing countries, an increase in energy use decreases income inequality. Therefore, any policy that reduces energy demand helps to enhance income inequality in developing countries and, thus, serves as a threat to the welfare level of the public by increasing deprivation and has a direct bearing on income inequality (Mahmood and Shah 2017).

3. Methodology, Variables, and Model

Income inequality (*IE*) is the dependent variable, energy crisis (*EC*) is an independent variable, and oil price shocks (*OPS*), gross domestic product (*GDP*), and population (*POP*) are controlled variables. Interaction terms are introduced to capture the indirect impact of other independent variables. The role of the energy crisis (EC) as a determinant of income inequality (IE) is grounded in the theoretical and empirical findings of Grunewald et al. (2017), Awan et al. (2013), and Mahmood and Shah (2017). These studies collectively establish that s scarcity of energy production is a fundamental cause of the energy crisis, which, in turn, exacerbates income inequality. The theory of inequality (Piketty 2014) is advanced to be employed for understanding inequality at the national level. This theory states that, if the returns on capital remain greater than economic growth, this could lead to *IE*. The household final consumption expenditures (*HFCE*) are negatively affected due to the increase in *EC*; as a result, households' capacity and ability to access energy supplies is reduced, causing inequalities. As such, Bazilian and Yumkella (2015) believe that a lack of access to energy increases economic deprivation and decreases business opportunities, and this deprivation translates into *IE*. Similar propagation is contemplated by Iddrisu and Bhattacharyya (2015), who theorized that energy deprivation is a vital source of *IE* in a society, if not justified based on a rational resource distribution mechanism. The following Table 1 contains the list of variables with units of measurement and sources.

Table 1. Variables' measurement and sources.

Variable	Unit of Measurement	Data Source
Energy Crisis	Kilos of Oil Equivalent	World Bank Data Bank, Pakistan Energy Year Books, Pakistan Economic Survey, Bureau of Statistics
Political Instability	Index (0–100)	International Country Risk Guide
Inflation Rate	Percentage	Pakistan Bureau of Statistics, Economic Survey of Pakistan
Population	Numbers	Pakistan Bureau of Statistics, Economic Survey of Pakistan
Income Inequality	Rate/Percentage	Pakistan Bureau of Statistics, Economic Survey of Pakistan, House Hold Integrated Economic Survey of Pakistan, Ministry of Finance
Gross Domestic Product	United State Dollars	Pakistan Bureau of Statistics, Economic Survey of Pakistan, Ministry of Finance
Household Final Consumption Expenditure	United State Dollars	Pakistan Bureau of Statistics, Economic Survey of Pakistan, House Hold Integrated Economic Survey of Pakistan, World Bank Data Bank
Oil Price Shocks	Net Oil Price Increase	World Bank Data Bank

The data utilized in this analysis present substantial benefits, such as the credibility of sources like the World Bank and Pakistan Bureau of Statistics, and the extensive inclusion of economic, social, and energy factors. This enables a thorough analysis over time with precise measurements. Nevertheless, it is important to take into account certain restrictions, as there may be discrepancies or inconsistencies between different sources of data, the Political Instability index is subjective, and there could be variations in inflation and oil price data across different regions.

3.1. Income Inequality

Income Inequality (*IE*) is a reflection of unequal resource distribution in society. That is, if the distribution of resources is biased toward one fraction of society compared to another fraction of the same society, it leads toward disparities, and the income gap between haves and have-nots, privileged and underprivileged, increases, thus resulting in *IE*. In this study, *IE* is measured using the theoretical proposition of Piketty (2014, 2015). He introduced a novel paradigm to address inequality, notwithstanding the traditional Gini coefficient, such that the return on capital is a critical factor in the determination of income distribution in society. The measure of *IE* states that, when the rate of return on capital (r) is higher than the economic growth rate (g) over a reasonable period, it leads toward *IE* in society; $r > g \rightarrow IE$. This measure of *IE* is denoted in Equation (1).

$$IE_t = r_t - g_t; \qquad t = 1, 2, \ldots, T \tag{1}$$

Therefore, this study developed a measure for determining *IE* based on the theories stated above, where the outcome of Equation (2) determine the situation of *IE* in society. That is, if $IE = 0$ means no or negligible income inequality and if $IE = +ve$ implies that resources are biased toward the capitalist fraction of a society, then the possibility of *IE* increases. On the other hand, if $IE = -ve$ implies that a society's share of national resources is increasing, *IE* is thus decreasing. According to Piketty (2014), the formula for calculating the rate of return from the capital is $r = \alpha\frac{g}{s}$, $\alpha = r.\beta$, and $\beta = \frac{s}{g}$.

The results of the DOLS regression analysis indicate that the energy crisis had a significant positive impact on income inequality. Table 1 presents the residual ADF test results, demonstrating the stationarity of the residuals. The Wald test results, shown in Table 2, confirm the existence of a long-term relationship between energy crisis and income inequality. The F-statistic (14.80) and chi-square statistic (88.83) are both significant at the 5% level, rejecting the null hypothesis of no long-run relationship. These findings support the third objective of this study and complement Piketty's hypothesis on resource distribution and inequality.

Table 2. VAR lag length selection criteria.

Lag	LogL	LR	FPE	AIC	SC	HQ
0	−2852.15	NA	1.95×10^{49}	136.19	136.52	136.31
1	−2469.66	601.04	5.33×10^{42}	121.03	124.01	122.12
2	−2317.29	181.40 *	1.11×10^{41}	116.08 *	122.45 *	118.88 *
3	−1967.67	125.15	7.25×10^{3}	126.83	122.58	118.93

* Indicates lag order selected by the criterion.

α = Share of capital income in national income;
r = Rate of return from capital;
β = Capital income ratio;
s = gross national savings rate;
g = economic growth rate.

Hence, with the changing scenarios in *EC*, the *IE* situation is hypothesized to change as well. Such that an increase in *EC* would inversely affect the buying abilities of household and provide fewer possibilities to access energy supplies. Therefore, the level of *IE* would

increase. Hence, it portrays a positive association between *EC* and *IE*. Since this research studies the effects of a higher rate of return on capital than the growth rate of the economy as a cause of income inequality in the case of Pakistan, the following variables are extracted.

3.2. Dynamic Ordinary Least Square Model

The dynamic ordinary least squares (DOLS) is used herein to achieve objective three of this study. It is a dynamic method for addressing the long-run time series of macroeconomic dynamic relationships. Particularly when those variables are cointegrated, it is imperative to understand long-run cointegrating vectors, as well as the extended cointegrated system in parameterizing in the form of an error correction mechanism; redefining the approach of DOLS became inevitable for a better and more simple explanation. Therein, the complexity of macroeconomic dynamics requires a better approach than to be handled by simple regression under OLS, especially in the presence of un-denied long-run relationships between variables in a time series and measuring the cointegrating vectors, with plausible endogeneity of the regressors. The DOLS, as a convincing solution, comes forward with a single-equation approach that addresses the endogeneity of a regressor while incorporating the leads and lags of the first difference in the regressor, and moreover, it addresses the issue of serially correlated error terms (Stock and Watson 1993). Masih and Masih (1996) had an extended application of DOLS to analyze the energy market dynamics of energy prices and demand. Therefore, it validated the role of this method as a reliable technique to inquire about the relationship between energy market dynamics and household capacities, abilities, and income inequality. A simpler version of Stock and Watson's (1993) regression can be generated using Equation (2), as follows:

$$y_t = \beta_0 + \beta_1 z_{1t} + \beta_2 z_{2t} + \beta_3 x_t + \varepsilon_t \quad (2)$$

Assume that y_t and z_t contain stochastic trends but x_t and ε_t do not. Here, only β_3 is asymptotically normal, whereas our hypothesis tests are not asymptotically valid for the variables that contain a stochastic trend (ε_t). The contribution of Stock and Watson (1993) changes Equation (4) so that the β_1 and β_2 coefficients apply to nontrending variables. Hence, DOLS comes forward as a method of estimating cointegrated models. After re-specifying and adjusting for no serial correlation, Equation (4) can be written as Equation (3), as follows:

$$y_t = \beta_0 + \beta_1 z_{1t} + \beta_2 z_{2t} + \beta_3 x_t + \beta_4 \Delta z_{1t} + \beta_5 \Delta z_{2t} + \varepsilon_t \quad (3)$$

where Δ represent change, and Δz_{1t} does not contain a stochastic trend. However, the paucity of serial correlation complicates DOLS; therefore, the model is extended in Equation (4), as follows:

$$y_t = \beta_0 + \beta_1 z_{1t} + \beta_3 x_t + \varepsilon_t \quad (4)$$

where y_t and z_{1t} contains a stochastic trend. Rewriting Equation (4) to reach to Equation (5) produces the following:

$$y_t = \beta_0 + \beta_1 z_{1t} + \beta_3 x_t + \gamma_1 \Delta z_{1,t+2} + \gamma_2 \Delta z_{1,t+1} + \gamma_3 \Delta z_{1,t} + \gamma_4 \Delta z_{1,t-1} + \gamma_5 \Delta z_{1,t-2} + \varepsilon_t \quad (5)$$

Therefore, a general, composite, and single equation for Stock and Watson could be summarized in Equation (6), as follows:

$$y_t = \alpha_0 + \beta X + \sum_{j=-q}^{p} d\Delta lny_{t-j} + \varepsilon_t \quad (6)$$

where α_0 is the intercept; t = 1,2, . . ., n; and j = 1,2, . . ., q, β is a co-integrating vector, representing a long-run cumulative multiplier. The purpose of including lag and lead lengths is to make the stochastic error term independent of past innovations in stochastic regression.

The study also proposes an integrated model for *IE* based on the theories and discussions in the preceding sections. Based on Stock and Watson (1993), the study developed a

DOLS model for understanding the relationship between *EC* and *IE*. The primary function declares that the *IE* is a function of *EC*, along with other controlled variables, provided that all variables are I(1); then, the study proceeds to implement DOLS in Equation (8). The long-run estimate of the optimal DOLS model is adjusted herein based on Stock and Watson (1993) and Masih and Masih (1996).

The links between energy access and reducing inequalities have been evident in the form of access to and the supply of energy. In Pakistan, while the majority may have access to energy supplies, underprivileged and middle-class households often spend a disproportionate portion of their incomes on energy due to higher prices. The nonavailability of low-cost energy supplies mainly affects underdeveloped regions with low-income households. Therefore, augmenting and extending supplies of modern, domestic, cleaner, indigenous, and low-cost energy is significant for this stratum of society. It is equally essential to reduce the excessive share of these groups' expenditures on energy supplies and services. Cautious deliberation regarding energy tariff structures is inevitable, and safety nets and social security programs for the poorest should be expanded. At this stage, targeting energy prices to fulfill the objectives of reducing polarities and increasing real incomes is crucial. Expanding cost-effective and efficient household devices related to housing, transportation, agriculture, small domestic scale production, and water pumping is also necessary. Energy policy initiatives should transform from a binary understanding to a qualitative paradigm regarding access, availability, provision of energy, and income inequality.

where $t = 1,2, \ldots, n$; $j = 1,2, \ldots, q$; and IE_t are dependent variables, EC_t is a matrix of the explanatory variable; and β is the co-integrating vector, which represents the long-run cumulative multiplier. The direct relationship of explanatory variables with the dependent variable is denoted by

where β_i, ($i = 1, 2, \ldots, 7$) and the indirect relationships between explanatory variables are represented through interaction terms, which are denoted by d_i, ($i = 1, 2, \ldots, 7$). In the end, the unit root test is performed on the residuals of the estimated DOLS regression in order to test for spurious regression. After the confirmation of the nonstationary residuals, the co-integration regression proceeds.

Diagnostic Checking

The prevailing literature on econometrics lays a strong emphasis on the procedural fairness and quality of a model's specification, such as the assumption of normality of model's error term, structural specification of the model, the choice of regressors, and measurement error. There is no doubt that diagnostic testing is now firmly established as a central topic in both econometric theory and practice (Hendry et al. 1980). DOLS by Stock and Watson (1993) has good robustness in the context of deviations from assumptions of standard regression, especially when it comes to residual correlation, heteroscedasticity, misspecification of functional form, and nonnormality of residuals. The stability tests were employed in the form of the CUSM and CUSMSQ so that the stability of the model could be verified and validated. Furthermore, one of the conditions for the application of DOLS is that the stationarity of the DOLS model's residual is checked in the long run and subsequently employed.

4. Results and Discussion

The results for the lag selection criteria, as shown in Table 1, are for the optimal lag for the DOLS, wherein the values of AIC, LR, FPE, and SC are all statistically significant. However, the value of AIC is 116.8, which is the lowest among all of the criteria and is significant at two lags.

The presence and significance of cointegration, which is a necessary condition for the application of DOLS, was tested and the results are displayed in Table 2.

The trace test specifies at least five cointegrating equations which have a *p*-value less than $\alpha = 0.05$, implying that the H0 of no cointegration is rejected. The trace results show evidence of having superior performance in cases in which the process has two

or more cointegrating relations than the amount specified under the null hypothesis. Lütkepohl (2001) concluded by advising that both tests should be executed simultaneously or that the trace test should be applied exclusively. The max-eigenvalue test shows that at least five cointegrating equations have *p*-values less than $\alpha = 0.05$; thus, the H0 of no cointegration is rejected. Therefore, the model satisfies the condition of the long-run relationship amongst variables. Henceforth, the proposed DOLS for dynamic long-run analysis is applied.

4.1. Long-Run Dynamic Results

The long-run DOLS estimation results are exhibited in Table 3, which shows the long-run relationship between *EC*, *PIS*, *OPS*, and other controlled variables, such as *GDP*, *INF*, and *HFCE,* and income inequality (*IE*). The model opts for one less lag than the prescribed optimal lag selection and, thus, relies on two leads. The presence of leads along with lag is to support in maintaining the asymptotic significance of the model's residual, which is a mandatory condition before the application of DOLS. Therefore, the test enables us to institute the possibility of a dynamic long-run affiliation in the model.

Table 3. Johansen's cointegration test.

No. of CE(s)	Eigenvalue	Statistic	Critical Value	Prob. **	Statistic	Critical Value	Prob. **
None *	0.93	110.17	52.36	0.00	347.73	159.52	0.00
At most 1 *	0.87	85.17	46.23	0.00	237.56	125.61	0.00
At most 2 *	0.74	56.50	40.07	0.00	152.38	95.75	0.00
At most 3 *	0.59	37.11	33.87	0.01	95.88	69.81	0.001
At most 4 *	0.52	30.54	27.58	0.02	58.77	47.85	0.003
At most 5	0.30	15.02	21.13	0.28	28.22	29.79	0.07
At most 6	0.16	7.51	14.26	0.42	13.19	15.49	0.10
At most 7 *	0.12	5.68	3.84	0.01	5.68	3.84	0.01

Trace test indicates 5 cointegrating equation(s) at the 0.05 level. Max-eigenvalue test indicates 5 cointegrating equation(s) at the 0.05 level. * Denotes rejection of the hypothesis at the 0.05 level. ** *p*-values.

The coefficient of *EC* is positive and statistically significant as the *p*-value is less than $\alpha = 0.05$. Thus, the H0 of there being no dynamic long-run relationship between *EC* and *IE* is rejected. This implies that a one KOE increase in the *EC* deteriorates the *IE* by one percent. This means that increasing the energy supply–demand gap is creating a resource distribution disparity, such that increasing *EC* implies that less energy supplies and services are available. Therefore, its distribution becomes biased and benefits the upper and middle classes more than the lower levels of the society, thus escalating the *IE*. This outcome also suggests that the existing energy infrastructure is more urban-centric. Therefore, the effect of a crisis falls more on the rural poor, due to which the growing turmoil enhances the *IE*. This outcome also strengthens the theory of resource distribution and inequality of Piketty (2015) while endorsing the third research objective of this study; this result is similar to Sharif Razavian et al. (2014), who showed the importance of enhancing energy use in development while testing DOLS for long-run dynamic relationships. The result is also in line with the path-dependency and uneven distribution of resource (energy) theory by Bouzarovski and Herrero (2017), who showed energy deprivation as the incapability of a household to secure a communally and substantially required level of energy services in the home, and when this situation is widespread across the country, its spatial and social distributions are highly uneven. Khan and Heinecker (2018) found different levels of energy services and supplies among different fractions of European society, which is particularly similar to findings by Mahmood and Shah (2017) in the case of Pakistan. Hence, they established that the *IE* can be explained by the current unjustified allocation of energy resources, in the long run, reflected in the persistent *EC*. In similar terms, deprivation of required energy is considered to be a social loss by Lutzenhiser (1993), with prior support and approval in the form of the capabilities theory by Sen (1992).

The negative and significant coefficient of *PIS* implies that improvements in political structures, processes, and sustainability reduce *IE*. The outcome also supports the theories of Lipset (1960) and Piketty (2014), which is an approval of this study's hypothesis. Therefore, the *PIS* is responsible for creating the situation of resource distribution disparity in the country. Since the disparities in the country do not seem to be improving despite maintaining an economic growth rate of 4.5 percent, this indicates, according to Easterly (2003), a weak political structure and system. As a weak institution cannot translate policy into implantation at the ground roots of the society, a fraction of society remains excluded and deprived of the resources distributed, and the presence of weak institutions and polity means that the extractive economic agents, as per Acemoglu and Robinson (2013), receive the disproportionate share and benefits, and therefore inequality increases.

The negative and significant impact of *OPS* on *IE* shows that increasing shocks reduce inequality. A possible explanation can be found in Malik et al. (2017), who showed the significant negative long-run impact of oil price shocks on economic growth, inflation, and interest rates. Moreover, Lakner (2016) opined that emerging economies portray greater results in terms of Piketty's hypothesis because of their low capital stock and high returns on capital and income growth are concentrated at the top. Thus, with Pakistan being a developing country with healthy returns on capital and heavily dependent on energy inputs from international markets (Malik 2012), the negative and statistically significant coefficient of *OPS*, therefore, reflects that the impact of oil price shocks is due to an increase in input costs and the cost of investments (Nazir and Hameed 2015). As a result, the return on capital starts to fall. Since the monetary policy, in response to the *OPS* tends to reduce interest rates to facilitate the business and capital accumulation process, if no change in interest rates or policy rates are assumed, then an increased *OPS* would adversely affect businesses. As a result, the rate of return on capital starts decreasing, which eventually decreases the level of *IE*.

Thus, the negative sign reflects that increasing oil prices and shocks affect the capitalist class more immediately than the households, and when the return on capital falls more than the adversity of the households, the inequality in terms of Piketty's model starts decreasing. As noted by Bourguignon et al. (2015), the capital has remained the main beneficiary of the globalization of trade and the resulting acceleration in the economic progress over the last two decades. Thus, the results show that as much as a one percent increase in the net oil price, considered oil price shocks as per Hamilton (1983) and Bernanke et al. (1997), could decrease income inequality by four percent. The significant part of measuring inequality using Piketty's model is that it reflects the inequality that occurred in society due to the extra flourishment of capital, enterprise, and the business. It also reflects that resources are being disproportionately devoted toward enterprise growth than toward development, as the country has a very meek indigenous production capacity and has massive reliance on oil imports, which account for almost 80 percent of its oil requirements. This result implies that the general price level is heavily dependent on oil price dynamics, and any volatility in the international oil prices would immediately affect the domestic energy sector and spread to all other sectors of the economy (Valasai et al. 2017).

The coefficient of *HFCE* is positive and statistically significant at the 10 percent level, implying that increasing the *HFCE* increases the *IE*. This reflects that an increasing rate of consumption expenditures exacerbates inequality in the country. As the measure of *IE* by Piketty (2014) incorporates the incomes and expenditures of the capitalist and rich fraction of the society (Lakner 2016), increasing expenditures therefore reflect more income and returns on capital. Thus, the gap keeps increasing with increasing household expenditure. Inflation has a positive and significant coefficient, implying that increasing inflation deteriorates the resource distribution fairness, and the lower-income fraction is more affected. As inflation in Pakistan is directly associated with world energy prices (Malik 2012), increasing prices, therefore, decreases the purchasing power of the household and keeps them deprived of energy services and supplies (Ali 2014). Moreover, Haider et al. (2014) believed that the inflation rate in Pakistan is predominantly driven by international energy

market dynamics and oil price shocks. Hence, the deprived fraction of the society loses purchasing power and the ability to deal with inflation. As a result, their actual incomes decrease, while inflation does not affect the upper fraction of the society. Resultantly, the gap increases, i.e., *IE*.

Moreover, the negative statistically significant coefficient of *GDP* implies that increased growth in the economy would reduce the pressure on *IE*. Therefore, the unstable economic situation of the country is a significant contributor to creating inequalities in resource distributions (Shahbaz 2010). Out of all the factors, terrorism and the failure of the government to control the threat are the most significant factors in the declining *GDP*. These adversely volatile conditions in the country scare off local, as well as foreign, investors from any existing or future investments. The decline in productive output and ever-increasing demand has created inflation, causing unemployment and poverty to prevail throughout these years (Tabassam et al. 2016). On the other hand, the positive and significant coefficient of *POP* implies that increasing the population is responsible for creating the relative deficiency in the resource supply, which creates the problem of *IE*. Hence, it can safely be inferred that increasing population levels over the years have been significant contributors to *IE* in the country. The distribution of authority and finances divides the community into different classes and results in an unstable and less fair distribution of resources. This imbalance gives opportunities to the elite for personal benefits. The misconduct and lack of administration lead citizens experience corruption at different levels. The decision-making policies ultimately benefit decisionmakers and the elite (Khan 2015).

4.2. Diagnostic Checking

There is a fair chance in econometric modeling for misspecification in regression analysis, with plausible adverse effects on the sampling properties of both the estimators and tests. Thus, they possess equal implications when attempting inferences and forecasting for the fitted model. The prevailing literature on econometrics lays strong emphasis on procedural fairness and the quality of the model's specification, such as the assumption of normality of the model's error term, structural specification of the model, the choice of regressors, and measurement error. Undoubtedly, diagnostic testing nowadays is recognized as a crucial subject in both econometric theory and practice. Some useful general references in this field include Krämer et al. (1986), Godfrey (1988, 1996), and White (1994), among many others. One of the conditions for the application of DOLS is that it requires us to check the stationarity of the DOLS model's residual in the long run. Hence, the estimations of DOLS in Table 4 proceed to Table 5 for the unit-root testing of its residual. The results show that the DOLS residual has no unit-root issue, as the *p*-value of ADF test is zero, which is less than $\alpha = 0.05$, thus rejecting the H0 of a unit root in the residual of DOLS model.

After the clarification of stationarity for the long-run DOLS, it is pertinent to address the presence of the long-run relationship through the Wald Test. The results are exhibited in Table 6. As *p*-values of $\chi 2$ and *F*-statistics is 0.00 and 0.01 (See Table 7), respectively, is less than $\alpha = 0.05$. Therefore, the H0 of a no long-run relationship is rejected. Hence, it implies the exists of a significant long-run relationship. The results thus support and affirm the third objective of this study and complement Piketty's hypothesis of resource distribution and inequality.

A schematic representation of the CUSUM is presented in Figure 2 and the CUSUMSQ in is presented Figure 3, establishing that the stability condition falls within the satisfactory section with the Akaike graph that underpinned the maximum lag selection criterion.

Table 4. Long-run dynamic estimation results.

Variable	Coefficient	Std. Error	*t*-Statistic	Prob.
EC	1.00	0.22	4.36	0.01 *
PIS	−8.40	2.04	−4.11	0.01 *
OPS	−4.08	1.62	−2.52	0.06 **
INF	6.11	1.68	3.63	0.02 *
HFCE	8.80	4.05	2.17	0.09 **
GDP	-3.19×10^{-9}	6.97×10^{-10}	−4.57	0.01 *
POP	3.11×10^{-6}	1.01×10^{-6}	3.08	0.03 *
C	−1882.1	413.29	−4.55	0.01 *
R-Squared	0.96	Mean Dependent Var.	11.41	
Adjusted R-Squared	0.68	S.D. Dependent Var.	20.45	
S.E. of Regression	11.46	Sum Squared Resid.	525.54	
Long-Run Variance	53.26			

*, ** Denote significance at 5 percent and 10 percent, respectively.

Table 5. Null hypothesis: D (DOLSRESIDUALS) has a unit root.

		t-Statistic	Prob.
Augmented Dickey–Fuller Test Statistic		−6.26	0.00 *
Test Critical Values:	1% level	−3.65	
	5% level	−2.95	
	10% level	−2.61	

* Denotes significance at the 5 percent.

Table 6. Dynamic DOLS residual ADF test.

Variable	Coefficient	Std. Error	*t*-Statistic	Prob.
D(DOLSRESIDUALS(-1))	−8.80	1.40	−6.26	0.00 **
D(DOLSRESIDUALS(-1),2)	6.47	1.28	5.04	0.00 **
D(DOLSRESIDUALS(-2),2)	5.21	1.07	4.84	0.001 **
D(DOLSRESIDUALS(-3),2)	3.91	0.85	4.57	0.001 **
D(DOLSRESIDUALS(-4),2)	2.65	0.63	4.20	0.003 **
D(DOLSRESIDUALS(-5),2)	1.48	0.40	3.69	0.001 **
D(DOLSRESIDUALS(-6),2)	0.50	0.16	3.07	0.005 *
C	0.15	0.58	0.27	0.788
R-Squared	0.91	Mean Dependent var.		−0.37
Adjusted R-Squared	0.89	S.D. Dependent var.		10.02
S.E. of Regression	3.31	Akaike Info Criterion		5.44
Sum Squared Resid.	263.63	Schwarz Criterion		5.81
Log-Likelihood	−79.14	Hannan–Quinn Criteria		5.56
F-Statistic	37.08	Durbin–Watson Stat.		2.02
Prob. (F-Statistic)	0.000			

*, ** Denote significance at the 5 percent and 10 percent, respectively.

Table 7. Wald Test for DOLS.

Test Statistic	Value	df	Prob.
F-statistic	14.80	(6, 4)	0.01 *
χ^2	88.83	6	0.00 *
Null Hypothesis: C(2) = C(3) = C(4) = C(5) = C(6) = C(7) = 0			

* Denotes significance at 5 percent.

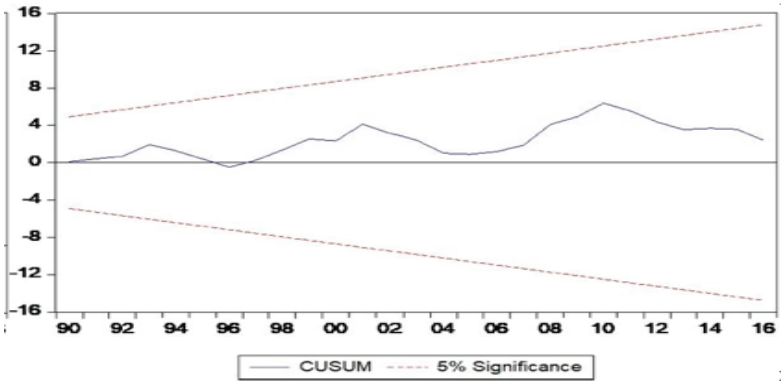


Figure 2. CUSUM.

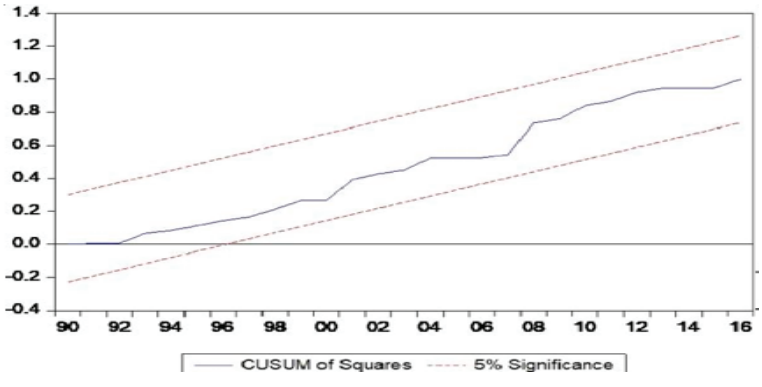


Figure 3. CUSUMSQ.

5. Conclusions and Policy Recommendations

The links between energy and reducing inequalities have been witnessed in an evident manner in the forms of the access to and the supply of energy. In Pakistan, the majority of people may have access to energy supplies. However, it is often the fractions of society that are underprivileged, below the extreme poverty line, and the middle classes who are using disproportionate amounts of their incomes to pay for energy supplies and services; to some extent, this is because of the higher upfront prices of energy supplies, expensive products, and expensive imported energy appliances. The nonavailability of low-cost energy supplies mainly affects underdeveloped regions, which have the majority of low-income households. However, this is equally applicable to middle- and lower-income groups of the country, who generally compose the majority living in underprivileged rural areas, for whom energy is an inelastic fundamental necessity of living, and a significant fraction of their energy consumption is to ensure basic survival. Therefore, augmenting and extending the supplies of modern, domestic, cleaner, indigenous, and low-cost types of energy are significant for this stratum of society. At the same time, the equally inevitable requirement is to diminish the excessive share of these groups' expenditures on energy supplies and services. The analysis of the results in the previous section shows that in most of the country, low-income groups spend a more significant share of their income on energy products, supplies, and services than higher-income fractions.

Fair and equal access to energy supplies and services is less likely to reduce income inequality if it is not cost efficient, which is an unfortunate fact in the case of Pakistan. The electricity tariff systems create nonproductive indications for the provision of low-cost

supplies of energy instead of leading and facilitating access to energy sources. Cautious deliberation regarding energy tariff structures is inevitable, but at the same time, the safety nets and social security programs for the poorest needed to be expanded. At this stage, the energy supply is targeting energy prices that will substantially fulfil the objectives of reducing polarities and increasing real incomes, whereas it is equally important to expand cost-effective and efficient household instruments related to housing, transportation, agriculture, small domestic-scale production processes, and water pumping. Initiatives on energy policy are required to transform from a binary understanding to a qualitative paradigm regarding access, availability, provision of energy, and income inequality. There is a need to quantify the actual fraction of the total population that is deprived of basic, as well as cost-effective, energy supplies and services, and, at the same time, to encompass the barriers to providing cheap energy supplies to rural areas and constraints of small- and medium-household investments, through the Scheme for Financing Renewable Projects (State Bank of Pakistan 2016), and the associated risks. Therefore, the first plan of action for the government should be to facilitate the private sector and maintain the ease of doing business by reducing security and the related investment risks so that small-household financial policy de-risking instruments may be introduced and technical services may be provided geared toward the installation of low-cost renewable energy products for the low-income households, as per the Alternative and Renewable Energy Policy (2019). Ideally, it should start in rural areas which have fewer opportunities to access modern energy services and supplies. The factors that influence universal access to energy services, supplies, and quality, as well as equity should be the fundamental objective of socioeconomic and sociopolitical decision-making institutions. The political elite should moderate their consensus around the provision of affordable energy supplies to reduce energy poverty, energy deprivation, and inequalities. This may require the implementation of strict transparency measures in the energy sector, harmonizing the provinces, including Gilgit-Baltistan, Azad Jammu Kashmir, private stakeholders, and provincial governments, for integrated energy policies for a grassroots-level transformation. The draft of an inclusive, unified, integrated energy policy to be enforced for private investors, public enterprises, and federal and local governments alike to expand and offer special energy services and supplies to rural areas and the underprivileged via clean and green initiatives should be developed. Meanwhile, regulators can devise a strategy to link energy tariffs and taxes to income levels and regulate the pricing accordingly.

Future research should focus on evaluating the impact of electricity pricing structures on different income groups and the effectiveness of social safety nets in reducing high electricity prices. It is also important to examine the barriers to the use of renewable energy solutions that benefit rural areas. In addition, research should examine how political ideology and understanding affect sustainable energy policy and explore how the private sector can contribute to enhancing public performance in improve energy access and affordability.

Author Contributions: Conceptualization, S.S.; data curation, R.A., Y.Y. and Q.A.; formal analysis, S.S. and J.H.; funding acquisition, L.J.; methodology, J.H., R.A. and S.S.; project administration, L.J.; resources, S.S. and Y.Y.; writing—original draft, S.S., Q.A. and R.A.; writing—review and editing, Y.Y., S.S. and Y.Y. All authors have read and agreed to the published version of the manuscript.

Funding: Ministry of Science and Technology, China. National Foreign Expert Project High-End Foreign Expert Introduction Plan. Educational traffic congestion management in China's megacities for the "carbon peaking and carbon neutrality" Strategy—Machine learning model based on National Big Data for New Energy Vehicles (NDANEA) (QN2022178002L).

Institutional Review Board Statement: Not Applicable.

Informed Consent Statement: Not Applicable.

Data Availability Statement: The data can be obtained upon request from the corresponding authors.

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

References

Acemoglu, Daron, and James Robinson. 2013. *Why Nations Fail: The Origins of Power, Prosperity, and Poverty*. New York: Crown Publishing Group.

Aguiar, Mark, and Mark Bils. 2015. Has consumption inequality mirrored income inequality? *American Economic Review* 105: 2725–56. [CrossRef]

Ali, Sharafat. 2014. Inflation, Income inequality, and economic growth in Pakistan: A cointegration analysis. *International Journal of Economic Practices and Theories* 4: 33–42.

Alternative and Renewable Energy Policy. 2019. Available online: https://climate-laws.org/document/alternative-and-renewable-energy-policy-2019_991c (accessed on 19 June 2023).

Alvaredo, Facundo, and Leonardo Gasparini. 2015. Recent trends in inequality and poverty in developing countries. In *Handbook of Income Distribution*. Edited by Anthony B. Atkinson and François Bourguignon. Amsterdam: Elsevier, vol. 2.

Arslan, Muhammad, Rashid Zaman, and R. K. Malik. 2014. Impact of CNG load shedding on daily routine: A study of Pakistan. *International Letters of Social and Humanistic Sciences* 42: 110–20. [CrossRef]

Atkinson, Anthony B., Thomas Piketty, and Emmanuel Saez. 2011. Top incomes in the long run of history. *Journal of Economic Literature* 49: 3–71. [CrossRef]

Awan, Rehmat Ullah, Falak Sher, and Akhtar Abbas. 2013. An investigation of multidimensional energy poverty in Pakistan. *The Pakistan Development Review* 4: 65–75.

Bazilian, M., and K. Yumkella. 2015. *Why Energy Poverty is the Real Energy Crisis*. Global Governance. Available online: https://www.weforum.org/agenda/2015/03/why-energy-poverty-is-the-real-energy-crisis (accessed on 19 June 2023).

Bernanke, Ben S., Mark Gertler, Mark Watson, Christopher A. Sims, and Benjamin M. Friedman. 1997. Systematic monetary policy and the effects of oil price shocks. *Brookings Papers on Economic Activity* 1997: 91–157. [CrossRef]

Berthe, Alexandre, and Luc Elie. 2015. Mechanisms explaining the impact of economic inequality on environmental deterioration. *Ecological Economics* 116: 191–200. [CrossRef]

Bouzarovski, Stefan, and Sergio Tirado Herrero. 2017. The energy divide: Integrating energy transitions, regional inequalities and poverty trends in the European Union. *European Urban and Regional Studies* 24: 69–86. [CrossRef]

Bourguignon, Thierry, Anne-Lorraine Bouquiaux-Stablo, Pascal Candolfi, Alain Mirza, Claudia Loardi, Marc-Antoine May, Rym El-Khoury, Michel Marchand, and Michel Aupart. 2015. Very long-term outcomes of the Carpentier-Edwards Perimount valve in aortic position. *The Annals of Thoracic Surgery* 99: 831–37. [CrossRef]

Boyce, James. 1994. Inequality as a cause of environmental degradation. *Ecological Economics* 11: 169–78. [CrossRef]

Checchi, Daniele, and Cecilia García-Peñalosa. 2008. Labour market institutions and income inequality. *Economic Policy* 23: 602–49. [CrossRef]

Cheema, Ahmed Raza, and Maqbool H. Sial. 2012. Poverty, income inequality, and growth in Pakistan: A pooled regression analysis. *The Lahore Journal of Economics* 17: 137–57. [CrossRef]

Ceyhan, Gaye D., and Deniz Saribas. 2022. Research trends on climate communication in the post-truth era. *Educational and Developmental Psychologist* 39: 5–16. [CrossRef]

Dalton, Hugh. 1920. The measurement of the inequality of incomes. *The Economic Journal* 30: 348–61. [CrossRef]

Easterly, William. 2003. The political economy of growth without development. In *In Search of Prosperity: Analytic Narratives on Economic Growth*. Princeton: Princeton University Press, pp. 439–72.

EIA, USA. 2019. Available online: https://www.eia.gov/energyexplained/energy-and-the-environment/where-greenhouse-gases-come-from.php (accessed on 19 June 2023).

Galvin, Ray, ed. 2019. *Inequality and Energy: How Extremes of Wealth and Poverty in High Income Countries Affect CO_2 Emissions and Access to Energy*. New York: Academic Press.

Godfrey, Leslie G. 1988. *Mis-Specification Tests in Econometrics: The Lagrange Multiplier Principle and Other Approaches*. Cambridge: Cambridge University Press.

Godfrey, Leslie G. 1996. Mis-specification tests and their uses in econometrics. *Journal of Statistical Planning and Inference* 49: 241–60. [CrossRef]

Grunewald, Nicole, Michael Jakob, and Ioanna Mouratiadou. 2014. Decomposing inequality in CO_2 emissions: The role of primary energy carriers and economic sectors. *Ecological Economics* 100: 183–94. [CrossRef]

Grunewald, Nicole, Stephan Klasen, Inmaculada Martínez-Zarzoso, and Chris Muris. 2017. The trade-off between income inequality and carbon dioxide emissions. *Ecological Economics* 142: 249–56. [CrossRef]

Haider, Adnan, Qazi Masood Ahmed, and Zohaib Jawed. 2014. Determinants of energy inflation in Pakistan: An empirical analysis. *The Pakistan Development Review* 53: 491–504. [CrossRef]

Hamilton, James D. 1983. Oil and the macroeconomy since World War II. *The Journal of Political Economy* 91: 228–48. [CrossRef]

Haq, Rashida, and Nasim Shah Shirazi. 1998. Trends in Inequality and Welfare in Consumption Expenditure: The Case of Pakistan. *The Pakistan Development Review* 37: 765–79. [CrossRef]

Heerink, Nico, Abay Mulatu, and Erwin Bulte. 2001. Income inequality and the environment: Aggregation bias in environmental Kuznets curves. *Ecological Economics* 38: 359–67. [CrossRef]

Hendry, Nico, Abay Mulatu, and Erwin Bulte. 1980. Econometrics: Alchemy or science? *Economica* 47: 387–406.

Iddrisu, Insah, and Subhes C. Bhattacharyya. 2015. Sustainable energy development index: A multi-dimensional indicator for measuring sustainable energy development. *Renewable and Sustainable Energy Reviews* 50: 513–30. [CrossRef]
International Atomic Energy Agency. 2005. *Energy Indicators for Sustainable Development: Guidelines and Methodologies*. United Nations, Department of Economic and Social Affairs, International Energy Agency. Vienna: International Atomic Energy Agency.
International Energy Agency. 2017. *Energy Access Outlook 2017*. Paris: OECD.
Keynes, John Maynard. 1973. *The Collected Writings of John Maynard Keynes. 7, The General Theory*. London: Macmillan.
Khan, Fouad, and Paul Heinecker. 2018. Inequality and energy: Revisiting the relationship between disparity of income distribution and energy use from a complex systems perspective. *Energy Research & Social Science* 42: 184–92.
Khan, Muhammad Arshad. 2015. Modelling and forecasting the demand for natural gas in Pakistan. *Renewable and Sustainable Energy Reviews* 49: 1145–59. [CrossRef]
Krämer, Walter, Harald Sonnberger, Walter Krämer, and Harald Sonnberger. 1986. Diagnostic checking in practice. *The Linear Regression Model Under Test*, 123–55.
Kuznet, Simon. 1955. Economic growth and income inequality. *The American Economic Review* 45: 1–28.
Kuznets, Simon. 1955. International differences in capital formation and financing. In *Capital Formation and Economic Growth*. Princeton: Princeton University Press, pp. 19–111.
Lakner, Christoph. 2016. *Global Inequality: The Implications of Thomas Piketty's Capital in the 21st Century*. World Bank Policy Research Working Paper, (7776). Washington, DC: The World Bank.
Liddle, Brantley. 2013. Impact of population, age structure, and urbanization on carbon emissions/energy consumption: Evidence from macro-level, cross-country analyses. *Springer Science, Business* 35: 286–304. [CrossRef]
Lipset, Seymour Martin. 1960. Party systems and the representation of social groups. *European Journal of Sociology/Archives Européennes de Sociologie* 1: 50–85. [CrossRef]
Lutzenhiser, Loren. 1993. Social and behavioral aspects of energy use. *Annual Review of Energy and the Environment* 18: 247–89. [CrossRef]
Lütkepohl. 2001. Vector autoregressions. In *Companion to Theoretical Econometrics', Blackwell Companions to Contemporary Economics*. Oxford: Basil Blackwell, pp. 678–99.
Lyubimov, Ivan. 2017. Income inequality revisited 60 years later: Piketty vs. Kuznets. *Russian Journal of Economics* 3: 42–53. [CrossRef]
Mahmood, Rafat, and Anwar Shah. 2017. Deprivation counts: An assessment of energy poverty in Pakistan. *The Lahore Journal of Economics* 22: 109–32. [CrossRef]
Malik, Afia. 2012. *Power Crisis in Pakistan: A Crisis in Governance?* No. 2012: 1. Islamabad: Pakistan Institute of Development Economics.
Malik, Kashif, Haram Ajmal, and Muhammad Umer Zahid. 2017. Oil price shock and its impact on the macroeconomic variables of Pakistan: A structural vector autoregressive approach. *International Journal of Energy Economics and Policy* 7: 83–92.
Marx, Karl, and Friedrich Engels. 2008. *The Communist Manifesto*, 3rd ed. New York: Pathfinder Press.
Masih, Abul MM, and Rumi Masih. 1996. Energy consumption, real income and temporal causality: Results from a multi-country study based on cointegration and error-correction modelling techniques. *Energy Economics* 18: 315–34. [CrossRef]
Murtaza, Ghulam, and Muhammad Zahir Faridi. 2015. Causality linkages among energy poverty, income inequality, income poverty and growth: A system dynamic modelling approach. *The Pakistan Development Review* 54: 407–25.
Nazir, Sidra, and T. Hameed. 2015. Impact of oil price and shocks on economic growth of Pakistan: Multivariate analysis (sectoral oil consumption). *Business and Economics Journal* 6: 1.
Nwosu, Emmanuel O., Obed Ojonta, and Anthony Orji. 2018. Household consumption expenditure and inequality: Evidence from Nigerian data. *International Journal of Development Issues* 17: 266–87. [CrossRef]
Pao, Hsiao-Tien, Yi-Ying Li, and Hsin-Chia Fu. 2014. Clean energy, non-clean energy, and economic growth in the MIST countries. *Energy Policy* 67: 932–42. [CrossRef]
Piketty, Thomas. 2014. *Capital in Twenty First Century*. London and Cambridge: The Belknap Press of Harvard University Press.
Piketty, Thomas. 2015. Capital in the twenty-first century. *The American Economic Review* 105: 48–53. [CrossRef]
Polinsky, A. Mitchell. 1974. Economic Analysis as a Potentially Defective Product: A Buyer's Guide to Posner's" Economic Analysis of Law". *Harvard Law Review* 87: 1655–81. [CrossRef]
Ravallion, Martin, Mark Heil, and Jyotsna Jalan. 2000. Carbon emissions and income inequality. *Oxford Economic Papers* 52: 651–69. [CrossRef]
Schroeter. 2013. *Time to Get Serious About the Energy Crisis*. Cologny: World Economic Forum, p. 2.
Sen, Amartya. 1992. *Inequality Re-Examined*. New York: Clarendon Press, p. 155.
Shahbaz, Muhammad. 2010. Income inequality-economic growth and non-linearity: A case of Pakistan. *International Journal of Social Economics* 37: 613–36. [CrossRef]
Sahir, Mukhtar H., and Arshad H. Qureshi. 2007. Specific concerns of Pakistan in the context of energy security issues and geopolitics of the region. *Energy Policy* 35: 2031–37. [CrossRef]
Shehzadi, Iram, Hafiz Muhammad Abubakar Siddique, and M. Tariq Majeed. 2019. Impact of political instability on economic growth, poverty and income inequality. *Pakistan Business Review* 20: 825–39.
Sharafat, A. L. I., Humayun Rashid, and Muhammad Aamir Khan. 2014. The role of small and medium enterprises and poverty in Pakistan: An empirical analysis. *Theoretical and Applied Economics* 18: 593.

Sharif Razavian, Ali, Hossein Azizpour, Josephine Sullivan, and Stefan Carlsson. 2014. CNN features off-the-shelf: An astounding baseline for recognition. Paper presented at the IEEE Conference on Computer Vision and Pattern Recognition Workshops, Columbus, OH, USA, June 23–28; pp. 806–13.

Simcock, Neil, Harriet Thomson, Saska Petrova, and Stefan Bouzarovski. 2017. *Energy Poverty and Vulnerability: A Global Perspective*. Hague and Abingdon: Routledge.

Skocpol, Theda. 1977. Wallerstein's world capitalist system: A theoretical and historical critique. *American Journal of Sociology* 82: 1075–90. [CrossRef]

Smith, Adam. 1776. *An Inquiry into the Nature and Causes of the Wealth of Nations*. Glasgow Edition of the Works and Correspondence of Adam Smith. Edited by Nicole Woolsey Biggart. New York: Oxford University Press, vol. II. Available online: http://www.ibiblio.org/ml/libri/s/SmithA_WealthNations_p.pdf (accessed on 19 June 2023).

State Bank of Pakistan. 2016. Annual Report 2016. Available online: https://www.sbp.org.pk/reports/annual/arFY16/Anul-index-eng-16.htm (accessed on 19 June 2023).

Stock, James H., and Mark W. Watson. 1993. A simple estimator of cointegrating vectors in higher order integrated systems. *Econometrica: Journal of the Econometric Society* 61: 783–820. [CrossRef]

Tabassam, Aftab Hussain, Shujahat Haider Hashmi, and Faiz Ur Rehman. 2016. Nexus between political instability and economic growth in Pakistan. *Procedia: Social and Behavioral Sciences* 230: 225–34. [CrossRef]

Toman, Michael A., and Barbora Jemelkova. 2003. Energy and economic development: An assessment of the state of knowledge. *The Energy Journal* 24: 93–112. [CrossRef]

Valasai, Gordhan Das, Muhammad Aslam Uqaili, HafeezUr Rahman Memon, and Saleem Raza Samoo. 2017. Overcoming electricity crisis in Pakistan: A review of sustainable electricity options. *Renewable and Sustainable Energy Reviews* 72: 734–45. [CrossRef]

York, Richard, Eugene A. Rosa, and Thomas Dietz. 2004. The ecological footprint intensity of national economies. *Journal of Industrial Ecology* 8: 139–54. [CrossRef]

White, Halbert. 1994. *Estimation, Inference and Specification Analysis*. Cambridge: University Press.

World Bank. 2013. *Toward a Sustainable Energy Future for All: Directions for the World Bank Group's Energy Sector*. World Bank's Group Energy Sector. Washington, DC: World Bank. Available online: http://documents.worldbank.org/curated/en/745601468160524040/pdf/795970SST0SecM00box377380B00PUBLIC0.pdf (accessed on 19 June 2023).

World Bank. 2018. Available online: https://documents.worldbank.org/en/publication/documents-reports/documentdetail/630671538158537244/the-world-bank-annual-report-2018 (accessed on 19 June 2023).

World Bank. 2022. Available online: https://www.worldbank.org/content/dam/meetings/external/annualmeeting-1/AM22-WBG-Publications-Catalog.pdf (accessed on 19 June 2023).

World Bank Group. 2014. World Bank. Available online: http://data.worldbank.org/indictor (accessed on 19 June 2023).

Zaidi, Syed. Akbar. 2015. *Issues in Pakistan's Economy*, 3rd ed. Oxford: Oxford University Press USA.

Zouhaier, Hadhek, and Kefi Mohamed Karim. 2012. Democracy, investment and economic growth. *International Journal of Economics and Financial Issues* 2: 233–40.

economies

MDPI

Article

Determinants of Income Inequality in South Africa: A Vector Error Correction Model Approach

Molepa Seabela, Kanayo Ogujiuba * and Maria Eggink

School of Development Studies, University of Mpumalanga, Mbombela 1200, South Africa
* Correspondence: kanayo.ogujiuba@ump.ac.za

Abstract: The issue of income disparity has long plagued South Africa because of the political environment that existed before the country's 1994 democratic transition. Based on the widely used Gini index, which gauges global inequality, the nation routinely has some of the highest rates of income disparity in the world. Income inequality in South Africa keeps rising even after a number of frameworks and policies have been put in place, which has a big influence on society. Thus, it is essential to comprehend the causes of income disparity and put suitable policies in place to remedy it. The purpose of this study is to look into the relationship between South Africa's income disparity and its determinants. Using the Vector Error Correction Model (VECM) approach, this study empirically examines the effects of government spending on social grants, gross savings, population growth, and economic growth on income inequality from 1975 to 2017. Data on the Gini index are sourced from the Standardized World Income Inequality Database (SWIID). Findings reveal a statistically significant negative correlation between government spending on social grants and income inequality. Moreover, income inequality demonstrates a negative relationship with both gross savings and economic growth. However, population growth exhibits a positive correlation with income inequality. This study highlights the significance of implementing a comprehensive strategy to address income inequality in South Africa. This strategy should involve augmenting government expenditure on social grants, cultivating a savings culture within households, and enacting policies that incentivize job creation, particularly in areas with rapid population growth. In addition to making a substantial contribution to the body of evidence already available on income disparity, this study offers insightful information to policymakers working to improve the socioeconomic climate in South Africa.

Keywords: income inequality; government spending on social grants; Gini coefficient; VECM; South Africa

JEL Classification: C51; D63; H53

Citation: Seabela, Molepa, Kanayo Ogujiuba, and Maria Eggink. 2024. Determinants of Income Inequality in South Africa: A Vector Error Correction Model Approach. *Economies* 12: 169. https://doi.org/10.3390/economies12070169

Academic Editor: Fabio Clementi

Received: 20 April 2024
Revised: 3 June 2024
Accepted: 5 June 2024
Published: 1 July 2024

1. Introduction

In recent years, there has been a growing global focus on income inequality, particularly its significant impact on developing countries such as South Africa. This study delves into the complex dynamics of income inequality within South Africa irrespective of a myriad of political and socioeconomic challenges confronting it. It examines the intricate relationship between government spending on social grants and income inequality. Using the Vector Error Correction Model (VECM), this study provides a detailed analysis of these dynamics.

Ever since the country's democratic transition in 1994, South Africa has faced growing anxiety over the problem of increasing wealth disparity. The persistent upward trend in income inequality highlights the complexity of addressing its underlying causes through policy and intervention methods, even in the face of widespread implementation of social spending measures. The political environment that South Africa finds itself in right now presents challenges to these endeavors. However, the available empirical academic research indicates conflicting relationships between the amount of money the government spends

on social grants and income disparity. Additionally, there are still unanswered questions about the relationships between gross savings, population increase, economic expansion, and income inequality.

To comprehend income inequality both between and within groups, various inequality measures come into play such as the Gini coefficient, a widely utilized measure ranging from zero (indicating perfect equality) to one (indicating perfect inequality). Drawing from the Standardized World Income Inequality Database (SWIID), this study examines Gini indices for income inequality, considering both gross and net incomes from 1975 to 2017. The South African government through the National Development Plan (NDP) aims to reduce income inequality from 0.70 to 0.60 by the year 2030. Despite the various drafted and implemented policies, South Africa remains the most unequal country globally (IMF 2020). Severe income inequality has persisted over the last century, indicating the necessity of comprehensive economic reforms either through policy or legal prescripts. As Leibbrandt and Shipp (2019) opined, reducing inequality requires targeted policies that address disparities in earnings and employment prospects and ongoing support to diminish gender and racial income inequalities.

Irrespective of the political climate and orientation held by those in authority, this study contends that existing policies and frameworks fall short in the face of growing income inequality, urging a re-evaluation of the contributing factors to income inequality. This study focuses on South Africa, a country known for its struggle with income inequality. It provides valuable insights that can inform national policy and global discourse on income inequality. This study's inclusion of data spanning from 1975 to 2017 allows for a comprehensive analysis of long-term trends and dynamics, providing a strong foundation for policy recommendations and future research directions. Considering the aforementioned background, this study investigated how government socioeconomic spending, particularly on social grants, affects income inequality. Additionally, the roles of gross savings, population growth, and economic growth in shaping income inequality were analyzed. The paper follows the following layout: Section 2 critically analyzes the literature, encompassing both the theoretical framework and empirical studies on the relationship between government spending on social grants and income inequality. This section delves into existing research to provide a comprehensive understanding of the topic. Section 3 outlines the data and methodology utilized in the study. Section 4 presents the findings and engages in a detailed discussion. Finally, Section 5 concludes the study, offering policy recommendations based on the results of the study.

2. Related Literature

Income inequality, a pressing issue in both developed and developing economies, is influenced by a variety of factors. This literature review consolidates insights from seminal works, global- and emerging-market studies, and specific studies focused on Africa, particularly South Africa, to comprehensively understand how each factor affects income inequality. These factors are not only crucial in the context of income inequality but are also connected to important works in economics that offer both theoretical foundations and empirical evidence for understanding the complex relationships between them and income inequality.

2.1. Seminal Works

Although exactly what causes income inequality cannot be determined due to the variety of factors that play a role, in this study some of the crucial factors are identified. Financial development, economic growth, external trade, and government initiatives can help mitigate income inequality, whereas inflation worsens it (Kapingura 2017).

A fundamental component of Keynesian economics, government expenditure on social welfare programs plays a crucial role in managing economic fluctuations and lessening income inequality (Vo et al. 2019). Such policies can stabilize aggregate demand, achieve full employment, and mitigate income inequality (Vo et al. 2019). In his influential work

during the Great Depression, John Maynard Keynes established the framework for Keynesian economics, highlighting the significance of aggregate demand in propelling economic activity (Jahan et al. 2014). Keynes proposed government intervention through expansionary fiscal policies to address economic fluctuations, attain full employment, stabilize prices, and diminish income inequality (Jahan et al. 2014; Alamanda 2020).

Fishburn and Willig (1984) extended Dalton's principle of transfer for income redistribution, demonstrating that socially desirable transfers, coupled with inverse transfers at higher income levels, yield positive social benefits. They linked these transfer principles to measures of income inequality and social welfare. Lerman and Yitzhaki (1995) developed a method to decompose changes in the Gini coefficient into components that narrow the income gap and reorganize income rankings. When this approach was used to analyze U.S. taxes and transfers for 1991, it was demonstrated that fiscal policies have the potential to lessen inequality by reducing income gaps and reordering income rankings.

Apart from government intervention in the form of redistribution of income, savings can also play a role in income inequality. The relationship between savings and income inequality is complex. While savings can contribute to wealth inequality, they can also serve as a safety net and reduce poverty (Vo et al. 2019). Shen and Zhao (2022) found that the impact of savings on income inequality varies across different subgroups and economies. Serven and Schmidt-Hebbel (1999) found no evidence that income inequality affects aggregate saving across countries, whereas a study by Schmidt-Hebbel and Serven (2000) highlighted the theoretical ambiguity in the relationship between income inequality and aggregate savings, with empirical results showing no systematic effect. The relative income hypothesis (RIH) according to Duesenberry (1949) and the permanent income hypothesis (PIH) proposed by Friedman (1957) are two important theories for understanding consumption patterns, indicating that consumption patterns are influenced by relative income and expected lifetime income, respectively. In modern societies, social status and relative income have a major impact on consumption patterns. A study by Bisset and Tenaw (2020) showed that low-income individuals adjusted their consumption to keep pace with others, showing strong "proof effects" and "ratchet effects". A study by Palley (2010) found that wealthy households save more permanent income than poor households, suggesting that relative income influences consumption patterns. Stable, long-term income policies can reduce income inequality by stabilizing consumption patterns (Yun et al. 2023).

Another component that is identified by the literature as playing a role in income inequality is population growth. High population growth is generally associated with a less equal income distribution (Ram 1984; Oyekale et al. 2004; Kaasa 2005). Reducing population growth tends to increase the income share of the poorest segments of the population (Rodgers 1983; Oyekale et al. 2004). Lower population growth and limited migration may contribute to increased national and global economic inequality (Peterson 2017).

Although economic growth may seem important for the reduction of income inequality, the relationship between income inequality and economic growth is more complex. Simon Kuznets (1955) proposed the Kuznets curve, which shows an inverted U-shaped relationship between economic development and income inequality. This suggests that inequality is a temporary phase in the development process. Arthur Lewis's (1954) dual-sector model explains economic development through labor transfer from a traditional to a modern sector, initially increasing inequality but eventually decreasing it as more workers transition to higher-paying industrial jobs (Sumner 2018). Lewis stressed the importance of government intervention to facilitate this transition and ensure fair income distribution (Sumner 2018). Piketty (2014) argues that returns on capital exceed the rate of economic growth and capital returns are higher than wages. He suggests that income inequality increases because wages grow more slowly than returns on capital. Piketty's work emphasizes the need for progressive taxation and policies to promote equal access to education and opportunities in order to address income inequality (Sawyer 2015). Mo (2000) developed a theoretical framework revealing that income inequality negatively influences GDP growth, particularly through the transfer channel. Empirical studies in-

dicate that while economic growth can reduce poverty, income inequality can intensify poverty and exacerbate the impact of growth on poverty (Amponsah et al. 2023). Economic growth exhibits poverty-reduction properties, but income inequality intensifies poverty and aggravates the impact of growth on poverty (Adeleye et al. 2020). The impact of GDP growth on poverty reduction diminishes with higher initial inequality, with a smaller poverty-reduction response in sub-Saharan Africa (Fosu 2009).

Evidence from empirical studies to determine how these theories apply and how these factors of income inequality play a role will be discussed further.

2.2. Africa/South African Studies

Leibbrandt et al. (2012) discovered that social transfers, particularly child support grants and old-age pensions, played a crucial role in decreasing poverty and income inequality in South Africa. Woolard et al. (2015) demonstrated that progressive taxes and pro-poor social spending significantly reduce income inequality in South Africa. The findings indicate a negative relationship between progressive taxes and pro-poor social spending. Additionally, Schiel et al. (2014) found that while social grants have helped alleviate poverty, they have not significantly reduced income inequality in South Africa. Household composition decomposition techniques revealed that changes have significantly reduced the direct impact on inequality through changes in household composition. Consequently, the relationship between government expenditures and income inequality is deemed insignificant. Despite the vital role played by social grants in reducing South Africa's persistently high levels of inequality, greater efforts are needed to further reduce income inequality as it remains relatively high.

The impact of savings on economic growth in South Africa is negative in the long run but positive in the short run (Van Wyk and Kapingura 2021). The development of the financial sector, especially when inclusive, can reduce income inequality, making financial inclusion crucial for benefiting disadvantaged groups (Kapingura 2017). According to Yun et al. (2023), the government may consider introducing a policy that allows tax deductions for retirement savings. Additionally, the government can design welfare and social security programs to provide individuals with a steady income over the long term rather than short-term cash payments. High population growth in low-income countries, including many in Africa, may slow economic development and exacerbate income inequality (Peterson 2017). Limited migration and lower population growth could increase economic inequality both nationally and globally. Studies by Nwosa (2019) and Ullah et al. (2021) support the positive relationship between population size and income inequality.

The relationship between economic growth and income inequality in African countries, including South Africa, has been the subject of various studies on the Kuznets curve. The Kuznets hypothesis, which suggests an inverted U-shape relationship between economic growth and income inequality, has been challenged in several empirical studies. These studies rejected the hypothesis because the data used were cross-sectional, meaning that the countries analyzed were at different stages of development. For example, Wahiba and Weriemmi (2014), Niyimbanira (2017), Nwosa (2019), Mdingi and Ho (2021), and Chude and Chude (2022) re-evaluated the relationship between income inequality and economic growth and found that variations in income distribution are more related to country-specific characteristics than data comparability issues. In South Africa, high income inequality has been shown to have a negative impact on long-term economic growth (Mdingi and Ho 2021). Niyimbanira (2017) found that in the Mpumalanga province of South Africa, economic growth was associated with poverty reduction but did not significantly affect income inequality, contrary to theoretical expectations. Chude and Chude (2022) and Nwosa (2019) found no significant effect of income inequality on economic growth in Nigeria. Wahiba and Weriemmi (2014) found that economic growth had a positive impact on income inequality in Tunisia. Zungu et al. (2021) discovered that lower growth is associated with lower income inequality in the SADC region. Despite efforts to address income inequality through economic growth and redistributive policies, South Africa

remains one of the most unequal countries globally (Francis and Webster 2019). According to studies by Bhorat et al. (2014), long-term economic growth in South Africa led to a decline in aggregate poverty but also increased inequality between 1995 and 2005. Nambie et al. (2023) found that financial inclusion and investment have a positive impact on economic growth, while income inequality and unemployment have a negative effect. Niyimbanira (2017) suggested that although economic growth can reduce poverty, it does not necessarily lead to a more equal distribution of income.

3. Data and Methodology

This study pursued exploring how different independent variables affect income inequality, using a meticulous methodology that included econometric analyses to gauge and analyze the impacts of crucial variables like government spending on social grants, gross savings, population growth, and economic growth. Importantly, the Gini coefficient index was used as a stand-in for income inequality, given its ability to offer a comprehensive gauge of income or wealth distribution within a population.

3.1. Data Source

This study used annual time series data between 1975 and 2017. The Gini coefficient is the metric used to measure income inequality, while government spending on social grants, gross savings, population growth, and economic growth serve as independent variables. The Gini index, which is the control variable, was obtained from the SWIID dataset by Solt (2020), while the independent variable data were sourced from the World Bank's World Development Indicators (WDI). The SWIID dataset was chosen due to its consistent time-series data and uniform data collection methods. This ensures comparability across countries throughout the specified study period.

3.2. Model Specification and Definition of the Variables

Model Specification:

The decision to use VECM in this context as the preferred model is supported by some previous studies. These include a study by Asari et al. (2011), who used VECM to examine the relationship between economic variables in emerging economies and demonstrated its effectiveness in capturing both short-term dynamics and long-term equilibrium relationships. Similarly, Arshad and Ali (2016) employed VECM in their study to investigate the relationship between the unemployment rate, interest rate, and inflation rate in Pakistan and analyze the models' short-term dynamics.

The variables are defined as follows:

$$Gini_t = \varnothing_0 + \varnothing_1 SG_t + \varnothing_2 GS_t + \varnothing_3 POPG_t + \varnothing_4 RGDPR_t + \varnothing_5 Dummy + \varepsilon_t \quad (1)$$

where:

- $Gini_t$ South Africa's Gini index, disposable income. The value assigned to this index is between 0 and 1, representing the dependent variable. A Gini index of zero indicates that there is no income inequality, while an index closer to one implies higher income inequality. Countries with a Gini index close to one are the most unequal in terms of income. This study focuses on the Gini index due to its widespread use, comparability, and availability over time, making it a suitable measure for analyzing and comparing income inequality trends. While the income quintile share ratio is informative, it is less familiar and less available in historical data compared to the Gini index.
- SG_t: Social grants include government spending on grants for the elderly, children, and disabled individuals, expressed as a percentage of the national budget. The coefficient is expected to be negative as government grants tend to reduce income inequality. While expressing SG as a percentage of GDP provides a broader economic view, the national budget perspective is particularly relevant in South Africa due to the country's history of social inequality and the crucial role of social grants in reducing

poverty and redistributing income. In addition, while disaggregating social transfers could provide valuable insights into their varying impacts on income inequality in South Africa, data limitations for the study period (1975–2017) prevent such analysis in the current study.

- ❖ GS_t: Gross savings represent the difference between disposable income and consumption and replace gross domestic savings, a concept used by the World Bank and included in World Development Indicators editions before 2006. Gross savings are calculated as gross national income minus total consumption plus net transfers. The anticipated coefficient for this variable is expected to be negative.
- ❖ $POPG_t$: Population growth (annual %) represents the total percentage change in population, assuming a constant growth rate between two points in time. The anticipated coefficient for this variable is expected to be negative.
- ❖ $RGDPR_t$: The annual growth of GDP at market prices, based on constant local currency and expressed in U.S. dollars, is calculated using aggregates based on constant 2015 prices. GDP encompasses the sum of gross value added by all resident producers, accounting for product taxes and subtracting subsidies not included in product values. This is a proxy for economic growth. The anticipated coefficient for this variable is expected to be negative.
- ❖ D_{ummy} = 1 from 2008 to the end of the sample and zero otherwise. This assumes it captures the period of the global financial crisis.
- ❖ ε_t: Represents the error term, encompassing other variables that may influence the relationship between the dependent variable and independent variables but were not explicitly included in the analysis.

3.3. Analytical Technique

The Vector Error Correction Model (VECM) was the chosen estimation technique in this study. As per the insights from Andrei and Andrei (2015), when a set of variables exhibits one or more cointegrating vectors, VECM becomes a suitable approach. The VECM is chosen for its ability to capture both short-term dynamics and long-term equilibrium relationships. This choice aligns with the objective of our study to explore the relationship between income inequality and its determinants in South Africa, which likely involves variables with cointegrating properties, as well as being particularly useful when dealing with cointegrated non-stationary time series data.

In this study, the Vector Error Correction Model (VECM) was deliberately chosen to conduct a comprehensive analysis of the relationship between income inequality and its determinants in South Africa. The primary objective was to account for both short-term fluctuations and long-term equilibrium adjustments, thus increasing the robustness and reliability of their findings. This study aims to contribute towards a deeper understanding of the dynamics of income distribution and provide a basis for evidence-based policy interventions to address the issue of income inequality in South Africa.

3.3.1. Unit Root Testing

This study employed the Augmented Dickey–Fuller (ADF) and Phillips–Perron (PP) tests to evaluate the stationarity of the variables. Conducting tests is essential to ensure the stationarity of time series variables before carrying out any econometric analysis to minimize the likelihood of obtaining spurious results. In addition, it aimed to ascertain if the variables fulfill the prerequisites of the VECM model, necessitating that variables are integrated into order one. Ensuring that variables are integrated as order one is crucial for the cointegration property, a fundamental aspect of the VECM. The unit root test equations for these tests are as follows:

$$\begin{aligned} Intercept\ and\ trend: \ \Delta Y_t &= \alpha + \beta T + \delta Y_{t-1} + \mu_t \\ Intercept: \ \Delta Y_t &= \alpha + \delta Y_{t-1} + \mu_t \\ None: \ \Delta Y_t &= \delta Y_{t-1} + \mu_t \end{aligned} \tag{2}$$

3.3.2. Testing for Cointegration

This study employed the Johansen cointegration approach (Johansen 1991) to examine the long-term relationship among the variables. Cointegration suggests that while individual variables may not be stationary, certain linear combinations of these variables may exhibit stationarity, indicating a lasting relationship. After confirming that at least one variable was integrated at order one I(1), the cointegration test was conducted using Johansen's (1991) maximum likelihood approach. This implies that variables X_t and Y_t are integrated at order one I(1) and exhibit a linear combination following regression.

To demonstrate cointegration, it is necessary to formulate the following equation:

$$\beth(r, r+1) = TIn(1 - \beth_{r+1}) \quad (3)$$

where:

$\beth(r, r+1)$ = likelihood ratio test statistic;

r = cointegration vectors;

T = sample size;

λ_r = estimated value for the i_{th} ordered eigenvalue from the π matrix.

4. Findings and Discussion

The findings include the descriptive statistics, the long- and short-run estimation and diagnostic tests which is then followed by a discussion of the findings.

4.1. Descriptive Statistics

The descriptive statistics for the variables under consideration are presented in Table 1, offering an organized and summarized perspective of the data for enhanced interpretability (Wooldridge 2019). The mean of the Gini is 2.341628 and the median is 2.010000. The data shows a large spread from the mean, indicated by a standard deviation of 0.827532. The results show that the Gini coefficient has a high positive skewness, and measures of centrality such as the mean and median show a skewness coefficient of 0.593320. In addition to that, the data show an increase in kurtosis to 1.968487, indicating that there is a gap in the distribution of Gini index values. In terms of the gross savings ratio, the data have a high kurtosis of 1.055535, indicating that there are outliers in the data and a reduced skewness due to population increase. The central tendency measures are significantly positively skewed, with a skewness of 4.521406.

Table 1. Summary of descriptive statistics (1975 to 2017).

South Africa	GINI	GS	POPG	RGDP	SG
Mean	2.341628	18.40556	1.894849	44,558.76	1.770333
Median	2.010000	17.00704	1.646040	43,910.25	1.888067
Maximum	3.910000	30.13734	3.497676	60,000.75	2.369546
Minimum	1.350000	13.49738	0.387278	28,061.25	1.111265
Std. Dev.	0.827532	4.521406	0.884247	9802.099	0.370982
Skewness	0.593320	1.055535	0.275616	−0.082303	−0.637557
Kurtosis	1.968487	2.948416	1.697031	1.625771	1.993379
Jarque–Bera	4.429237	7.989539	3.586171	3.432119	4.728569
Probability	0.109195	0.018412	0.166446	0.179773	0.094017
Sum	100.6900	791.4391	81.47851	1,916,027	76.12430
Sum Sq. Dev.	28.76199	858.6107	32.83947	4.04×10^9	5.780374
Observations	43	43	43	43	43

Our study employed the Gini coefficient from the SWIID database to assess inequality in pre-tax, pre-transfer (market) income. The Gini coefficient typically ranges from 0 to 1. However, our analysis converted it into an unbounded measure using the following formula: Gini/(100 − Gini). Following the approach outlined in Ahmad (2017), we then converted the unbounded measure to natural log values. Consequently, these log-transformed values of median, mean, and mode exceed 1.

4.2. Unit Root Tests: ADF Unit Root Test and Lag–Length Selection Criteria

The ADF unit root tests in Table 2 indicate that the variables are not stationary at level I(0), as evidenced by negligible *p*-values (*p*-value > 0.1, 0.05, and 0.01). However, at I(1), significant *p*-values (*p*-value < 0.01 for LSG and *p*-value < 0.1 for all other variables) suggest that all variables are stationary at first difference without trend. The Johansen cointegration test is employed to explore the long-term relationship between variables given their stationary nature at I(1).

Table 2. Unit root of all variables.

Variable	Levels (ADF Test) (*p*-Value in Brackets)	First Difference (ADF Test) (*p*-Value in Brackets)	Result
LGINI	−2.942736 (0.1610)	−4.958705 *** (0.0013)	I(1)
LSG	−1.786786(0.6933)	−3.336181 * (0.0775)	I(1)
LGS	−1.821231 (0.6765)	−6.554354 *** (0.0000)	I(1)
LPOPG	−1.369638 (0.1557)	−2.261930 ** (0.0247)	I(1)
LRGDP	−1.951057 (0.6098)	−4.674045 *** (0.0000)	I(1)

Note: (*): The rejection of the null hypothesis of unit root at the 10% significance level. (**): The rejection of the null hypothesis of unit root at the 5% significance level. (***): The rejection of the null hypothesis of unit root at the 1% significance level.

The above test (Table 2) aims to identify at least one cointegrating equation, with optimal lags selected for the subsequent Johansen cointegration test. This suggests that, through first differencing, all variables exhibit stationarity, as indicated by ADF values surpassing their critical values at a 5% significance level. This outcome provides a pathway for conducting a cointegration analysis.

4.3. Johansen Cointegration Test

To assess if the variables in this study have a long-term relationship, there was a need to conduct a Johansen cointegration test since these variables are stable at I(1) level. A failure of the test to identify cointegration would suggest the absence of a long-term association among the variables.

Table 3 presents the Johansen cointegration findings. The *p*-values of the Trace and Max-Eigen test statistics are significant at the 5% significance level, and both test statistics are greater than their respective critical values. As a result, it can be concluded that there is only one cointegrating equation. The conclusion follows that there is a long-term link between the variables and income inequality in South Africa. Therefore, the null hypothesis is rejected.

In Table 4, we used the Phillips–Perron (PP) test to determine if the five different variables, LGINI, LGS, LSG, LPOPG, and LRGDP, were stationary or not. The test results showed that all five variables were non-stationary in their original form but became stationary after the first differencing. This is denoted as I(1), indicating that each variable is integrated as order one.

Table 3. Johansen cointegration results.

Hypothesized No. of ce(s)	Trace Test			Maximum Eigen Test		
	Trace Statistic	**t-Critical Values**	***p*-Value**	**Max-Eigen Statistic**	**t-Critical Values**	***p*-Value**
None *	133.4235	69.81889	0.0000 *	58.99659	33.87687	0.0000
At most 1 *	74.42690	47.85613	0.0000 *	47.97227	27.58434	0.0000
At most 2	26.45463	29.79707	0.1157	18.51302	21.13162	0.1118

Note: * denotes the rejection of the null hypothesis at the 0.05 level.

Table 4. Phillips–Perron (PP) test results.

Variable	Levels (PP Test) (*p*-Value in Brackets)	First Difference (PP Test) (*p*-Value in Brackets)	Result
LGINI	−1.712052 (0.7282)	−5.085249 *** (0.0009)	I(1)
LGS	−1.831453 (0.6715)	−6.583596 *** (0.0000)	I(1)
LSG	−1.785250 (0.6941)	−8.358888 *** (0.0000)	I(1)
LPOPG	−1.269205 (0.8819)	−6.034506 *** (0.0001)	I(1)
LRGDP	−1.567222 (0.7890)	−4.479040 *** (0.0048)	I(1)

(***): Shows significance at the 1% significance level.

4.4. Long-Run and Short-Run Estimation Results of the VECM Model (1975 to 2017)

The subsequent step involved employing the Vector Error Correction Model (VECM) to analyze both short-term and long-term relationships given the established existence of long-term relationships between the variables. VECM accommodates short-term adjustments while constraining the long-term behavior of endogenous variables to converge to their cointegrating relationships. As a suitable model for measuring corrections from past disequilibrium, VECM is deemed necessary. The presence of a stable long-term association is indicated by a negative and significant coefficient in the VECM, suggesting that any short-term fluctuations between variables result in a steady long-term relationship.

Table 5 illustrates the long-term relationship between the study variables. The cointegrating Equation (1) demonstrates the long-term relationships between income inequality (LGINI) and the explanatory variables (LSG, LGS, LPOPG, and LRGDP). Negative coefficients signify a positive long-term association between the dependent variable and its explanatory factors, while positive coefficients indicate a negative long-term relationship. The cointegration Equation (1) reveals a negative long-term link between the Gini coefficient and government spending on social grants, gross savings, population growth, and annual GDP growth.

The significant error correction term, which ranges from zero to negative values, signifies a stable long-term equilibrium. In this study, the error correction term of −0.005705 suggests a stable cointegration relationship, indicating a 0.57% adjustment rate. This adjustment rate reflects the speed at which shocks to independent variables return the Gini coefficient to its equilibrium, reducing the income inequality gap by 0.57% in the short term. The R-squared of 0.542628 measures the explanatory power of the model; that is, 54% of the variation in the dependent variable is explained by independent variables. From the long-run equation of the VECM, the dummy(−1) variable captures the global financial crisis negatively and significantly leads to income inequality in South Africa.

Table 5. VECM results.

Cointegrating Equation	Cointegration Equation (1)					
LGINI(−1)	1.000000					
LGS(−1)	−2.725854					
	(0.60893)					
	[−4.47649]					
LPOPG(−1)	−2.938554					
	(0.31013)					
	[−9.47518]					
LRGDP(−1)	−6.987395					
	(1.40785)					
	[−4.96318]					
LSG(−1)	0.218470					
	(0.37415)					
	[0.58391]					
DUMMY(−1)	1.051549					
	(0.10666)					
	[9.85877]					
C	82.98360					
Error Correction:	**D(LGINI)**	**D(LGS)**	**D(LPOPG)**	**D(LRGDP)**	**D(LSG)**	**D(DUMMY)**
Cointegration Equation (1)	**−0.005705**	**0.095592**	**0.368434**	**0.012212**	**−0.030995**	**0.072930**
	(0.01862)	**(0.05053)**	**(0.05560)**	**(0.00884)**	**(0.06916)**	**(0.13498)**
	[−0.30642]	**[1.89174]**	**[6.62693]**	**[1.38085]**	**[−0.44817]**	**[0.54029]**

4.5. Discussion of the Results

The results indicate a significant negative impact of government spending on social grants on income inequality, with statistical significance at the 1% level. This suggests that an increase in such spending is expected to reduce income inequality in the long term, reflecting a negative relationship between the two variables. These findings align with existing studies and theoretical frameworks, such as the Keynesian theory, which underscores the state's role in mitigating income inequality through government expenditures and taxes. Similarly, the permanent income hypothesis and the relative income hypothesis propose that social security programs' government expenditure can help alleviate income inequality. However, the Kuznets hypothesis contradicts the Keynesian theory by suggesting that inequality initially rises in economic development's early phases but falls in later stages, a trend not supported by this study's results. South Africa is considered an upper-middle-income country that has a diverse economy and a significant population of over 62 million, according to the Census 2022 (Stats SA 2023). Although the country has advanced from its early developmental stages, it does not follow the Kuznets theory. According to the theory, income inequality usually rises during the early stages of economic development and then declines. However, this is not the case in South Africa, where income inequality continues to increase despite the country's economic growth. The outcomes of this study are consistent with prior empirical studies by Sanchez and Perez-Corral (2018), Anderson et al. (2017), Woolard et al. (2015), and Leibbrandt et al. (2012). Thus, the statistical analysis in this study rejects the null hypothesis of no relationship between government spending on social grants and income inequality. This study provides solid evidence for the efficacy of such policies, which could shape budgetary priorities and social welfare strategies.

Gross savings has a significant positive relationship with income inequality, with a 1% significance level. This means that when households save more, it leads to an increase in income inequality. Darku (2014) found that increased income inequality results in increased consumption by individuals in all income groups, leading to declining personal savings rates. According to Palley (2010), wealthy households save a higher percentage of their permanent income than poor households, leading to disproportionate wealth accumulation and investment returns and exacerbating income inequality. Similarly, empirical studies by Maaboudi et al. (2023) and Tran et al. (2020) found a positive relationship between gross savings and income inequality. However, studies by Van Wyk and Kapingura (2021), Yildirim (2020), and Deniz and Ozturkler (2010) found a negative relationship, while Halim et al. (2016) found no significant association between gross savings and income inequality. The theoretical studies that support these findings are Yun et al. (2023), Friedman (1957), and Duesenberry (1949). When applying the Relative Income Hypothesis (Duesenberry 1949) to this study, it is found that individuals with lower incomes might save less to keep up with the consumption patterns of wealthier individuals, leading to lower wealth accumulation and higher income inequality over time. On the other hand, the Permanent Income Hypothesis (Friedman 1957) suggests that wealthier households save more, leading to faster wealth accumulation than that of poorer households, resulting in increasing income inequality. Yun et al. (2023) suggest that policy intervention aimed at stabilizing income can reduce income inequality by stabilizing consumption patterns.

The findings indicated that the population growth coefficient at 1% was statistically significant. As a result, population growth had a negative impact on income inequality over time. This suggests that as the population increases, income inequality is likely to also increase. The potential positive impact of population growth could be attributed to the fact that if state resources do not increase in line with the population, the allocation for social programs, healthcare, and education may result in fewer resources per person. This could strain society and ultimately lead to a rise in income inequality. Therefore, these results are consistent with other studies by Ullah et al. (2021), Nwosa (2019), Peterson (2017), and Anyanwu (2016), which found that population growth in the long-term leads to increases in income inequality. In an ideal scenario, an increasing population can lead to more entrepreneurial activity, job creation, and overall economic growth. However, in South Africa, several factors complicate these relationships. Issues such as structural inequality, limited access to education and resources, regulatory barriers, and economic instability can hinder entrepreneurial efforts and business growth. Historical differences in South Africa, including the impact of apartheid, have resulted in persistent economic inequality that disproportionately affects marginalized communities. High unemployment rates, skill shortages, and inadequate infrastructure further impede the development of entrepreneurs and economic expansion. Addressing these fundamental problems is crucial to creating a conducive environment for entrepreneurship and business growth while harnessing the potential benefits of a larger population. While social subsidies are essential in addressing immediate challenges of poverty and income inequality, long-term sustainable solutions must focus on encouraging inclusive economic growth and empowering individuals and communities to participate meaningfully in the economy. To achieve this, a comprehensive approach is needed to address both the supply-side constraints facing businesses and the broader socio-economic factors that contribute to inequality and exclusion. Prioritizing policies that support entrepreneurs and economic empowerment can help South Africa create a more equitable and prosperous society for all citizens. Policymakers may need to re-evaluate population-related policies like immigration, family planning, and resource allocation strategies. Addressing the potential negative impact of population growth on income distribution necessitates a holistic approach that accounts for demographic trends and social policy frameworks.

This study's findings show that annual GDP growth rates have a positive impact on income inequality in the long term. The commonly held view that economic growth automatically leads to improved income distribution is not always true in most developing

countries, including South Africa. This indicates that the relationship between economic growth and income inequality is more complex and can vary depending on various factors, such as policy interventions and labor market dynamics. This is consistent with a study by Wahiba and Weriemmi (2014) that also found a positive relationship. Empirical studies by Mdingi and Ho (2021), Nambie et al. (2023), Jianu et al. (2021), Vo et al. (2019), Royuela et al. (2019), and Caraballo et al. (2017) found a negative relationship between economic growth and income inequality. Chude and Chude (2022), Nwosa (2019), and Niyimbanira (2017) found no significant effect of income inequality on economic growth contrary to theoretical expectations.

The dummy variable representing economic crises shows a negative relationship with income inequality. This means that in the long-run equation of the VECM, the dummy variable has a negative and significant impact on income inequality in South Africa, particularly capturing the global financial crisis. This study's significant error correction term of −0.063277, which falls between zero and negative, indicates a stable long-run equilibrium. The negative error correction term also suggests a stable and statistically significant cointegration relationship. Policymakers may need to reconsider the connection between economic growth and income distribution, focusing on inclusive growth strategies that prioritize equitable wealth distribution. This finding underscores the importance of targeted interventions to ensure that economic prosperity benefits all members of society.

4.6. Robustness Check

Performing diagnostic tests is an essential part of this study since it indicates whether or not there is an issue with the model's estimation. If an issue is found, it indicates that the model is inefficient, which may also imply that the findings are skewed (Wooldridge 2001). Tests for normality, heteroscedasticity, and serial correlation were among the diagnostic procedures carried out to determine whether the model utilized in this investigation reasonably fits the data. The results of the diagnostic tests conducted for this study indicate that the model is quite well described. Table 6 shows that the residuals have a combined probability for the Jarque–Bera of 0.1107 and are normally distributed. The likelihood of 0.5015 for LM-Stat indicates that the residuals are not serially correlated. Furthermore, no heteroskedasticity has been discovered, as shown by a joint Chi-square probability of 0.5603.

Table 6. Diagnostic tests.

Test	Null Hypothesis	t-Statistics	Probability
Jarque–Bera (JB)	There is a normal distribution	4.401542	0.1107
Langrage Multiplier (LM)	No serial correlation	45.54572	0.5015
White (CH-sq)	No conditional heteroskedasticity	33.09478	0.5603

5. Conclusion and Recommendations

This study concludes by summarizing findings, providing recommendations, and outlining limitations in Sections 5.1–5.3.

5.1. Conclusions

South Africa has been struggling with the issue of income inequality for a long time, even before the onset of democracy. According to the widely accepted measure of global inequality, the Gini index, South Africa has the highest income inequality in the world. Despite implementing policies and frameworks, South Africa has seen a rise in income inequality, which has a profound impact on society. The purpose of this study was to analyze the relationship between specific economic indicators in South Africa using the Vector Error Correction Model (VECM) on income inequality. This is fundamental because

assessing the factors contributing to income inequality assists in finding appropriate measures to mitigate it. By evaluating government spending on social grants, gross savings, population growth, economic growth, and the dummy variable to capture economic crises, this study sheds light on their impact on income inequality from 1975 to 2017 in the South African context.

This study revealed that government spending on social grants has a negative impact on income inequality. This means that as government spending on social grants increases, income inequality is expected to decrease in the long term. Additionally, this study found that gross savings have a positive impact on income inequality, with a significance of 1%. This suggests that wealthier households tend to save a higher percentage of their permanent income compared to poorer households, leading to disproportionate wealth accumulation and exacerbating income inequality. Furthermore, this study's results indicated that economic growth has a positive impact on income inequality in the long run. However, it is worth noting that economic growth does not always lead to improved income distribution in developing countries like South Africa, as the relationship is influenced by factors such as policy interventions and labor market dynamics. Moreover, population growth is statistically significant at 1% and positively impacts income inequality in the long term. This implies that an increase in population over time can lead to a surge in income inequality. However, it is important to recognize that population growth can also stimulate entrepreneurship, create job opportunities, and contribute to overall economic development. In the long run, the dummy variable representing economic crises demonstrates a negative and significant relationship with income inequality in South Africa. This suggests that it has a notable impact on income inequality, particularly in the context of capturing the global financial crisis.

The results of this study add to the existing literature on the relationship between government spending, gross savings, population growth, economic growth, and income inequality. The findings highlight the importance of government spending on social grants and the negative impact of gross savings on income inequality. These findings can inform policy decisions to reduce income inequality in South Africa. To tackle the pressing problem of income inequality, policymakers are advised to adopt a versatile approach that includes policies aimed at economic growth and equitable income distribution. Such measures may include boosting government spending on social welfare programs and revamping social security policies, both of which have shown to be effective in addressing income inequality. In a South African context, a prime example was the successful implementation of the social relief of distress grant during the COVID-19 pandemic. Although increasing spending on social welfare programs can help to reduce poverty and income inequality in the short term, there are some potential disadvantages to consider. For instance, relying too heavily on social grants without implementing measures to promote economic growth and employment could lead to dependency instead of encouraging self-reliance. In addition, inefficient administration and corruption can undermine the effectiveness of social welfare programs, resulting in misallocated resources and worsening income inequalities.

This study's findings emphasize the significance of savings and economic growth in addressing income inequality. Encouraging people to save and invest can lead to economic growth and create opportunities for accumulating wealth. Policies that incentivize saving behavior, such as tax breaks and tax-free investment incentives, can effectively promote these endeavors without adversely affecting government income. Furthermore, implementing policies to strengthen economic growth, such as infrastructure development, innovation incentives, and trade facilitation, can stimulate job creation and income generation. Supporting small and medium-sized enterprises (SMEs) and entrepreneurship can also foster inclusive economic growth.

In addition, investing in population programs, including family planning initiatives and reproductive health services, can effectively manage population growth, therefore leading to less government spending on social grants. By empowering individuals to make informed choices about family planning, these programs can positively impact income

inequality by encouraging smaller family sizes. This, in turn, reduces pressure on resources and promotes economic development.

5.2. Recommendations

This study's findings highlight the importance of balancing individual saving behaviors with broader socio-economic goals, which can potentially influence financial regulation and social welfare policies.

The following recommendations are made based on the results of this study:

1. This study proposes implementing strategies to curb income inequality, including increasing government spending on social welfare programs and reforming social security policies. These measures have been demonstrated to be effective in mitigating income disparities, as exemplified by the success of the social relief of distress grant implemented during the COVID-19 pandemic in South Africa.
2. Policymakers are encouraged to address the fundamental causes of income inequality, acknowledging the essential role of labor supply and job creation in alleviating income inequality; policies should focus on employment expansion. Initiatives such as skill development programs can enhance the workforce's employability.
3. To balance population growth with inclusive economic development, policymakers are encouraged to develop policies that stimulate job creation and economic opportunities in regions experiencing rapid population growth. This should also foster an environment conducive to entrepreneurship and small business development to absorb the growing workforce and minimize the exacerbation of income inequality over the long term.
4. This study highlights the importance of policies geared towards improving gross savings. Encouraging a culture of saving and implementing incentives for individuals and businesses by the government can contribute to economic stability and resilience in the long term.

5.3. Limitations of This Study and Recommendation for Future Studies

This study had some limitations due to a shortage of relevant data and materials. This study is restricted to a specific time frame because of the availability of data. The study period is from 1975 to 2017, which means that there is a five-year time lag in terms of the data since this study concluded in 2023. The unemployment data were not included in the model due to the methodological constraints of the Vector Error Correction Model (VECM) and data limitations. The unemployment rate data are considered to be integrated as order zero (I(0)), indicating that it is already stationary. On the other hand, for the Vector Error Correction Model (VECM) to establish cointegrating relationships, all variables need to be integrated as order one (I(1)). If an I(0) variable is included in a VECM, it can result in model misspecification and unreliable results. The near singular matrix error occurred when the lagged Gini coefficient was introduced, indicating perfect collinearity. This means that the lagged Gini coefficient was highly correlated with the current Gini coefficient, making it redundant in the model. Including perfectly collinear variables violates the assumptions of the VECM and can lead to unstable estimates. Furthermore, despite several efforts that were made to collect relevant data from different sources, this study is constrained due to the limited amount of data on income inequality in South Africa. It is recommended that:

- Future research studies should investigate whether the results of this study would vary if the income inequality data were available over a more extended period.
- Also, future research could incorporate different categories of social transfers as separate variables, enabling a more nuanced examination of their effects on income inequality.

Author Contributions: Conceptualization, M.S. and K.O.; methodology, M.S.; validation, M.E., K.O. and M.S.; formal analysis, M.S.; investigation, M.S. and M.E.; data curation, M.E.; writing—original draft preparation, M.S. and K.O.; writing—review and editing, M.S., K.O. and M.E.; visualization, M.E.; supervision, K.O. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding, and the APC was funded by the University of Mpumalanga, South Africa.

Informed Consent Statement: Not applicable.

Data Availability Statement: Available on request.

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

References

Adeleye, Bosede Ngozi, Obindah Gershon, Adeyemi Ogundipe, Oluwarotimi Owolabi, Ifeoluwa Ogunrinola, and Oluwasogo Adediran. 2020. Comparative investigation of the growth-poverty-inequality trilemma in Sub-Saharan Africa and Latin American and Caribbean Countries. *Heliyon* 6: e05631. [CrossRef]

Ahmad, Mahyudin. 2017. Economic freedom and income inequality: Does political regime matter? *Economies* 5: 18. [CrossRef]

Alamanda, Alamanda. 2020. The effect of government expenditure on income inequality and poverty in Indonesia. *Journal Info Artha* 4: 1–11. [CrossRef]

Amponsah, Mary, Frank W. Agbola, and Amir Mahmood. 2023. The relationship between poverty, income inequality and inclusive growth in Sub-Saharan Africa. *Economic Modelling* 126: 106415. [CrossRef]

Anderson, Edward, Maria Ana Jalles D'Orey, Maren Duvendack, and Lucio Esposito. 2017. Does government spending affect income inequality? A meta-regression analysis. *Journal of Economic Surveys* 31: 961–87. [CrossRef]

Andrei, Dalina Maria, and Liviu C. Andrei. 2015. Vector error correction model in explaining the association of some macroeconomic variables in Romania. *Procedia Economics and Finance* 22: 568–76. [CrossRef]

Anyanwu, John C. 2016. Empirical Analysis of the Main Drivers of Income Inequality in Southern Africa. *Annals of Economics & Finance* 17: 337–64.

Arshad, Sumera, and Amajd Ali. 2016. Trade-off between Inflation, Interest and Unemployment Rate of Pakistan: Revisited. *Bulletin of Business and Economics (BBE)* 5: 193–209.

Asari, F. F. A. H., Nurul Syuhada Baharuddin, Nurmadihah Jusoh, Zuraida Mohamad, Norazidah Shamsudin, and Kamaruzaman Jusoff. 2011. A vector error correction model (VECM) approach in explaining the relationship between interest rate and inflation towards exchange rate volatility in Malaysia. *World Applied Sciences Journal* 12: 49–56.

Bhorat, Haroon, David Tseng, and Benjamin Stanwix. 2014. Pro-poor growth and social protection in South Africa: Exploring the interactions. *Development Southern Africa* 31: 219–40. [CrossRef]

Bisset, Tazeb, and Dagmawe Tenaw. 2020. Keeping up with the Joneses: The relevance of Duesenberry's relative income hypothesis in Ethiopia. *Research Square*. [CrossRef]

Caraballo, Mª Ángeles, Carlos Dabús, and Fernando Delbianco. 2017. Income inequality and economic growth revisited. A note. *Journal of International Development* 29: 1025–29. [CrossRef]

Chude, Nkiru Patricia, and Daniel Izuchukwu Chude. 2022. Effect of Income Inequality on the Economic Growth of Nigeria. *International Journal of Scientific and Management Research* 1: 22–35. [CrossRef]

Darku, Alexander. 2014. Income inequality, status seeking, and savings rates in Canada. *Canadian Studies in Population* 41: 88–104. [CrossRef]

Deniz, Pinar, and Harun Ozturkler. 2010. Effect of Income Inequality on Aggregate Saving: A Panel Analysis. Paper presented at Turgut Özal International Conference on Economics and Politics–I: Global Crises and Economic Governance, Malatya, Turkey, May 15–16; Available online: https://studys.ssrn.com/sol3/studys.cfm?abstract_id=3299364 (accessed on 10 February 2022).

Duesenberry, James Stemble. 1949. *Income, Savings, and the Theory of Consumer Behavior*. Cambridge: Harvard University Press.

Fishburn, Peter C., and Robert D. Willig. 1984. Transfer principles in income redistribution. *Journal of Public Economics* 25: 323–28. [CrossRef]

Fosu, Augustin Kwasi. 2009. Inequality and the impact of growth on poverty: Comparative evidence for Sub-Saharan Africa. *Social Science Research Network*. [CrossRef]

Francis, David, and Edward Webster. 2019. Inequality in South Africa. *Development Southern Africa* 36: 733–34. [CrossRef]

Friedman, Milton. 1957. The permanent income hypothesis. In *A Theory of the Consumption Function*. Princeton: Princeton University Press.

Halim, Mohammed Abdul, Sadia Sultana, Asma Akter, and Ataur Rahaman. 2016. Savings and Income Inequality in Bangladesh (Working Study Series: WP No 1617). Available online: https://www.bb.org.bd/pub/research/workingstudy/wp1617.pdf (accessed on 15 March 2020).

International Monetary Fund. 2020. Six Charts Explain South Africa's Inequality. Available online: https://www.imf.org/en/News/Articles/2020/01/29/na012820six-charts-on-south-africas-persistent-and-multi-faceted-inequality (accessed on 27 March 2024).

Jahan, Sarwat, Ahmed Saber Mahmud, and Chris Papageorgiou. 2014. Back to Basics: What Is Keynesian Economics?—The central tenet of this school of thought is that government intervention can stabilize the economy. *Finance & Development* 51: 53–54.

Jianu, Ionuț, Marin Dinu, Dragoș Huru, and Alexandru Bodislav. 2021. Examining the Relationship between Income Inequality and Growth from the Perspective of EU Member States' Stage of Development. *Sustainability* 13: 5204. [CrossRef]

Johansen, Søren. 1991. Estimation and Hypothesis Testing of Cointegration Vectors in Gaussian Vector Autoregressive Models. *Econometrica* 59: 1551–80. [CrossRef]
Kaasa, Anneli. 2005. Factors of Income Inequality and their Influence Mechanisms: A Theoretical Overview. *Social Science Research Network*. [CrossRef]
Kapingura, Forget Mingiri. 2017. Financial sector development and income inequality in South Africa. *African Journal of Economic and Management Studies* 8: 420–32. [CrossRef]
Kuznets, Simon. 1955. Economic growth and income inequality. *The American Economic Review* 45: 1–28.
Leibbrandt, Murray, Arden Finn, and Ingrid Woolard. 2012. Describing and decomposing post-apartheid income inequality in South Africa. *Development Southern Africa* 29: 19–34. [CrossRef]
Leibbrandt, Murray, and Timothy Shipp. 2019. Reducing Inequalities in South Africa: Progress on Equality Thwarted by Slow Growth and Success of Top Earners. (UN-WIDER Policy Brief 2019/2). Available online: https://www.wider.unu.edu/publication/reducing-inequalities-south-africa (accessed on 7 July 2022).
Lerman, Robert I., and Shlomo Yitzhaki. 1995. Changing Ranks and The Inequality Impacts of Taxes and Transfers. *National Tax Journal* 48: 45–59. [CrossRef]
Lewis, W. Arthur. 1954. Economic Development with Unlimited Supply of Labor. *The Manchester School of Economic and Social Studies* 22: 139–91. [CrossRef]
Maaboudi, Reza, Mohammad Hassan Fotros, and Zeynab Dare Nazari. 2023. The Effect of Income Inequality on the National Savings in Iran. *Stable Economy Journal* 4: 69–91.
Mdingi, Kholeka, and Sin-Yu Ho. 2021. Literature review on income inequality and economic growth. *MethodsX* 8: 1–12. [CrossRef] [PubMed]
Mo, Pak Hung. 2000. Income inequality and economic growth. *Kyklos* 53: 293–315. [CrossRef]
Nambie, Nicholas Bamegne, Philomena Dadzie, and Dorcas Oye Haywood-Dadzie. 2023. Measuring the Effect of Income Inequality, Financial Inclusion, Investment, and Unemployment, on Economic Growth in Africa: A Moderating Role of Digital Financial Technology. *International Journal of Economics and Financial Issues* 13: 111–24. [CrossRef]
Niyimbanira, Ferdinand. 2017. Analysis of the impact of economic growth on income inequality and poverty in South Africa: The case of Mpumalanga Province. *International Journal of Economics and Financial Issues* 7: 254–61.
Nwosa, Philip Ifeakachukwu. 2019. Income inequality and economic growth in Nigeria: Implication for Economic Development. *Acta Universitatis Danubius Oconomica* 15: 108–16.
Oyekale, A. S., A. I. Adeoti, and T. O. Ogunnupe. 2004. Sources of income inequality and poverty in rural and urban Nigeria. Paper presented at 3rd Annual Workshop of Poverty and Economic Policy (PEP) Network, Dakar, Senegal, June 11–20.
Palley, Thomas I. 2010. The Relative Permanent Income Theory of Consumption: A Synthetic Keynes–Duesenberry–Friedman Model. *Review of Political Economy* 22: 41–56. [CrossRef]
Peterson, E. Wesley F. 2017. The role of population in economic growth. *SAGE Open* 7: 215824401773609. [CrossRef]
Piketty, Thomas. 2014. *Capital in the Twenty-First Century*. Cambridge: Harvard University Press.
Ram, Rati. 1984. Population increase, economic growth, educational inequality, and income distribution. *Journal of Development Economics* 14: 419–28. [CrossRef]
Rodgers, Gerry. 1983. Population growth, inequality and poverty. *International Labour Review* 122: 443–60.
Royuela, Vicente, Paolo Veneri, and Raul Ramos. 2019. The short-run relationship between inequality and growth: Evidence from OECD regions during the Great Recession. *Regional Studies* 53: 574–86. [CrossRef]
Sanchez, A. Angeles, and Antonio L. Perez-Corral. 2018. Government social expenditure and income inequalities in the European Union. *Hacienda Pública Española* 227: 133–56. [CrossRef]
Sawyer, Malcolm. 2015. Confronting inequality: Review article on Thomas Piketty on Capital in the 2st Century. *International Review of Applied Economics* 29: 878–889. [CrossRef]
Schiel, Reinhard, Murray Leibbrandt, and David Lam. 2014. Assessing the Impact of Social Grants on Inequality: A South African Case Study. (UN-WIDER Working Study 2014/160). Available online: https://www.wider.unu.edu/publication/assessing-impact-social-grants-inequality (accessed on 24 June 2019).
Schmidt-Hebbel, Klaus, and Luis Serven. 2000. Does income inequality raise aggregate saving? *Journal of Development Economics* 61: 417–46. [CrossRef]
Serven, Luis, and Klaus Schmidt-Hebbel. 1999. *Income Inequality and Aggregate Saving: The Cross-Country Evidence*. World Bank Policy Research Working Paper. Washington, DC: World Bank. [CrossRef]
Shen, Chengfang, and Xindong Zhao. 2022. How does income inequality affects economic growth at different income levels? *Ekonomska Istraživanja/Ekonomska IstražIvanja* 36: 864–84. [CrossRef]
Solt, Frederick. 2020. Measuring income inequality across countries and over time: The Standardized World Income Inequality Database. *Social Science Quarterly* 101: 1183–99. [CrossRef]
Stats SA (Statistics South Africa). 2023. Census 2022 (Statistical Release P0301.4). Available online: https://census.statssa.gov.za/assets/documents/2022/P03014_Census_2022_Statistical_Release.pdf (accessed on 28 March 2024).
Sumner, Andy. 2018. Trickle along Not Trickle Down? Inequality and the Lewis Model of Economic Development. (ESRC GPID Research Network Briefing Paper 13). Available online: https://gpid.univie.ac.at/wp-content/uploads/2018/04/Brief_13.pdf (accessed on 22 May 2024).

Tran, Nhan Dang, Chuong Nguyen Ong, and Quy Dinh Le Nguyen. 2020. The relationship between income inequality and savings: Evidence from household-level panel data in Vietnam. *Journal of Applied Economics* 23: 709–28. [CrossRef]
Ullah, Atta, Zhao Kui, Saif Ullah, Chen Pinglu, and Saba Khan. 2021. Sustainable utilization of financial and institutional resources in reducing income inequality and poverty. *Sustainability* 13: 1038. [CrossRef]
Van Wyk, Bianca Flavia, and Forget Mingiri Kapingura. 2021. Understanding the nexus between savings and economic growth: A South African context. *Development Southern Africa* 38: 828–44. [CrossRef]
Vo, Duc Hong, Thang Cong Nguyen, Ngoc Phu Tran, and Anh The Vo. 2019. What Factors Affect Income Inequality and Economic Growth in Middle-Income Countries? *Journal of Risk and Financial Management* 12: 40. [CrossRef]
Wahiba, Nasfi Fkili, and Malek El Weriemmi. 2014. The Relationship between Economic Growth and Income Inequality. *International Journal of Economics and Financial Issues* 4: 135–43.
Woolard, Ingrid, Rebecca Metz, Gabriela Inchauste, Nora Lustig, Mashekwa Maboshe, and Catriona Purfield. 2015. How Much Is Inequality Reduced by Progressive Taxation and Government Spending. Available online: http://www.econ3x3.org/study/how-much-inequality-reduced-progressive-taxation-andgovernment-spending (accessed on 27 March 2024).
Wooldridge, Jeffrey M. 2019. Correlated random effects models with unbalanced panels. *Journal of Econometrics* 211: 137–50. [CrossRef]
Wooldridge, Jeffrey M. 2001. Diagnostic testing. In *A Companion to Theoretical Econometrics*. Edited by Badi H. Baltagi. Hoboken: Blackwell Publishing Ltd., pp. 180–200.
Yildirim, Ahmet Eren. 2020. The Relationships between Income Inequality & Savings: A Study on Households in Turkey. *BİLTÜRK Journal of Economics and Related Studies* 2: 370–80.
Yun, Wong Sing, Siti Hajar Samsu, Jain Yassin, Shairil Izwan Taasim, and Faerozh Madli. 2023. Household Consumption Function Based on Permanent Income Hypothesis: A Systematic Review. *International Journal of Academic Research Economics and Management Sciences* 12: 441–52.
Zungu, Lindokuhle Talent, Lorraine Greyling, and Nkanyiso Mbatha. 2021. Economic growth and income inequality: A non-linear econometrics analysis of the SADC region, 1990–2015. *African Journal of Economic and Management Studies* 12: 285–301.

economies

MDPI

Article

The Timing and Strength of Inequality Concerns in the UK Public Debate: Google Trends, Elections and the Macroeconomy

Knut Lehre Seip [1,*] and Frode Eika Sandnes [2]

[1] Faculty of Technology, Art and Design, Oslo Metropolitan University, P.O. Box 4, St. Olavs Plass, Pilestredet Park 33, 0176 Oslo, Norway
[2] Department of Computer Science, Oslo Metropolitan University, 0167 Oslo, Norway; frodes@oslomet.no
* Correspondence: knut.lehre.seip@oslomet.no

Abstract: Inequality among people has several unwanted effects, in humanistic, social and economic contexts. Several studies address distributional preferences among groups, but little is known about when inequality issues are focused and when and why inequality abatement measures are brought on the political agenda. We show that during the period 2004 to 2023, inequality issues were focused during elections to the EU and UK parliament and with greatest strength during the elections to the EU parliament in May 2004 and to the UK parliament in May 2015. Periods with high unemployment and inflation cause the discussion on inequality to be followed by discussions on inequality measures. However, when the discussion of inequality is followed very closely by the discussions of abatement measures, inequality aversion becomes more strongly associated with the macroeconomic variables inflation and GDP (recessions) than with unemployment and more strongly associated with the concerns for fairness than concerns with war and crime. The results were obtained examining Google Trends and scholarly studies.

Keywords: inequality; abatement measures; United Kingdom; Google Trends; parliament elections; recessions; unemployment; inflation

Citation: Seip, Knut Lehre, and Frode Eika Sandnes. 2024. The Timing and Strength of Inequality Concerns in the UK Public Debate: Google Trends, Elections and the Macroeconomy. *Economies* 12: 135. https://doi.org/10.3390/economies12060135

Academic Editor: Fabio Clementi

Received: 3 April 2024
Revised: 2 May 2024
Accepted: 21 May 2024
Published: 30 May 2024

1. Introduction

Inequality (INE) among people within countries and worldwide has been hypothesized to cause political instability, e.g., Roe and Siegel (2011), decrease economic growth, e.g., Michalek and Vybostok (2019), and contrast with what people regard as morally right, Rözer et al. (2022). Peoples' perception of inequality of pay has also been found to be strongly connected to their political standpoint (Bratanova et al. 2016; Fisman et al. 2017; Sandnes et al. 2023). However, Hoy et al. (2024, data from Australia) show that the opinions of right-leaning voters change when misconceptions are corrected. A popular survey of relations between inequality and population "spirit" is given in Wilkinson and Pickett (2010, data sources pp. 280–83). Three sets of questions with regard inequality can be distinguished: A first set addresses the factors that contribute to increasing or decreasing inequality. A second set addresses which of these factors are the result of economic policy to stabilize the economy, and which factors can be managed directly to address inequality, e.g., Malla and Pathranarakul (2022). The third set addresses the type of inequality concerns expressed by low-, middle- and high-income populations. Most of the current literature on inequality appears to address the concerns expressed by the low-income population.

Recently, Google Trends have been used to identify people's sentiments in the public conversation about important issues. For example, Timoneda and Wibbels (2022, pp. 3–6) discussed the use of Google Trends and examined the frequency of "protests" using Google Trends. Connor et al. (2019) searched for links between "inequality" and "racial-bias" but found no conclusive evidence for such a link. Here, we examine if the two issues, "inequality" (INE) and "inequality measures" (IMEs), were related to particular events or

macroeconomic economic states in the United Kingdom (UK). The UK was chosen because it ranks among the highest in Europe in terms of income inequality (Berisha et al. 2021), and Filippin and Nunziata (2019, p. 116) include the UK among the high-inequality cluster of European states.

We developed three hypotheses.

Hypotheses. We hypothesized that

Hypothesis 1 (H1). *Conversations about INE and measures to reduce inequality, IME, would be particularly strong around elections to the UK and the European union (EU) parliaments. The raationale is that when inequality becomes a focused issue in the public conversation, the need for measures to abate inequality will be stronger and therefore follow in strength in the conversation. For example, Fisman et al. (2017) argues that redistribution versus anticipated GDP growth was an important issue during the US presidential election in 2012. We evaluate the hypothesis by identifying time windows in Google Trends series where frequencies are larger than one half the series' average frequency; in the following, these are denoted by INE+ and IME+. The technique of identifying windows in time series where values that are higher and lower than the average value is common in climate research; e.g., Gong et al. (2020).*

Second, we hypothesized that

Hypothesis 2 (H2). *INE would be leading IME during unfavorable economic conditions. Such periods could be periods with high unemployment (UE), high inflation (INF) and low central bank interest rate (CBI). CBI is low when the economy is underperforming. We evaluate this hypothesis by (i) applying a high-resolution lead-lag method (HRLL) to paired Google time series and (ii) embedding the LL relations in a principal component analysis (PCA) "map" of the UK economy. This allows us to identify which economic conditions are associated with a leading relation for INE to IME, that is, from inequality discussions to inequality abatement measures.*

Third, we hypothesized that

Hypothesis 3 (H3). *A major reason for conflicts related to inequality would be "fairness" issues. We evaluated this hypothesis by comparing Google Trends series for "inequality" with Trends series for "fairness". If the two series should correlate, fairness is an important attribute of inequality.*

Since the political events associated with a parliament election last for a relatively short period, whereas economic changes typically last for the length of economic business cycle times, e.g., 4–8 years (Burns and Mitchell 1946), we disentangle the time series into a high-frequency, rapid component and a low-frequency, slow component.

We find that inequality in the UK is in focus during parliament elections, and during periods with recessions and inflations, more than with periods of unemployment. Inequality is more strongly associated with fairness than with crime and war.

The present study distinguishes itself from most other studies in that it identifies time windows during the economic and social development of a single country, the UK, whereas most other studies find patterns in inequality by studying and comparing inequality among several countries. We use a high-resolution lead-lag method (HRLL) that allows us to strengthen the causal interferences over what correlation alone would support. HRLLs are calculated over short time periods (3 months and 9 months are sufficient to establish significance). Furthermore, we add context to our results by making a PCA score plot for the UK economy from 2004 to 2023 based on five macroeconomic variables.

The rest of the manuscript is organized as follows. Section 2 presents the time series used. Section 3 describes the methods, with emphasis on a high-resolution lead-lag (HRLL) method. Section 4 shows the results, and Section 5 discusses the results. Section 6 presents concluding remarks.

2. Data

We used Google Trends https://trends.google.com/trends/?geo=NO, accessed on 5 January 2024. From that database, we retrieved time series for the frequency of the terms "inequality", "poor" and "rich" for the period 2004 to 2024. We used the colloquial terms "poor" and "rich" instead of low- and high-income people because the second set of terms did not have enough data to construct meaningful time series. We retrieved time series for "conflict", "political conflict" and "inequality measures" to express possible results of inequality on conflicts and abatement measures. We used the expression "Inequality measures" despite its relatively low overall scores because we then could examine if "inequality" would lead to the discussion of "inequality measures". We also tried several other expressions for inequality measures, but none yielded enough data to construct meaningful time series. To put our results into an economic context for the UK, we used data on unemployment (UE), monetary supply (M1), inflation (INF) and industrial production (IP). The data were extracted from St. Louis Federal Reserve in 5 January 2024 (https://fred.stlouisfed.org/). Data for UK and EU parliament elections were downloaded from official governmental websites. There were eight election dates: October 2004 (EU election), May 2005 (UK election), June 2009 (EU-election) 2010 May (UK election), May 2015 (UK election), June 2017 (UK election), May 2019 (EU election), December 2019 (UK election). The UK left the EU on 31 January 2020. We retrieved an index for inequality, the Gini's index from https://fred.stlouisfed.org/series/SIPOVGINIGBR and from https://www.statista.com/statistics/872472/gini-index-of-the-united-kingdom/, the latter to find the most recent scores. Both data sources were retrieved on 5 January 2024. The Gini index is an economic equality score calculated for all countries in the world by the World Bank (https://data.worldbank.org/indicator/SI.POV.GINI, accessed on 5 January 2024). Data for concerns related to inequality were retrieved from Google Trends, Google Scholar and Web of Science.

3. Methodology

In this section, we first disentangle the Google Trends time series into short-term and long-term components based on power spectral density (PSD) results for the Google series. We then explain how the series are detrended and smoothed to extract the features of interest. Thereafter we explain briefly the high-resolution lead-lag HRLL method used in the study, cycle periods (CP) and phase shifts (PS). Fourth, we briefly explain how we use some algorithms that are common in several statistical software packages.

3.1. Data Preprocessing

We disentangled the time series by identifying time series' components using LOESS smoothing. The LOESS smoothing algorithm has two parameters. The parameter (f) is the fraction of the time series that is used as a moving window, and the parameter (p) is the order of the polynomial equation used to interpolate the series. Because we always use p = 2, we use the nomenclature LOESS(f) to identify the smoothing degree in the rest of the manuscript. We LOESS(0.4)-smoothed the raw time series to identify the low-frequency time series and subtracted the low-frequency time series from the raw time series to identify the high-frequency time series. To remove high-frequency noise from the high-frequency series, we, in addition, used LOESS(0.2) to smooth the series.

3.2. Scoring Time Series Events

To calculate scores, we identified the number of months during the two recessions and during the eight elections where (i) "inequality" and (ii) "inequality measures" showed high frequency and (iii) where "inequality" was leading "inequality measures". Recessions lasted 16 and 6 months, respectively, and the election periods were defined as 2 months before and 1 month after the election month. This is the same time window as Timoneda and Wibbels (2022) use before and after a "protest" month.

3.3. Lead-Lag Analysis

We use the high-resolution lead-lag method to examine LL relations between the time series for "inequality" and "inequality measures". The description of the method closely follows the description given in Seip and McNown (2007). The method rests on the duality between paired time series as time series (x-axis as time, and the two series INE and IME on the y-axis) and as phase plots, with one series on the x-axis, INE, and the other series on the y-axis, IME. A persistent rotational direction in the phase plot will show that one series leads the other series in the time representation. The method allows us to select the degree of persistence of LL relations and to determine how closely the last series should follow the first series.

The method calculates an angle, θ(3), for two paired series $x(t)$ and $y(t)$ based on three consecutive paired observations in the phase plot for the series ($x = x(t)$ $y = y(t)$). Figure 1a shows two sine series with equal cycle periods (λ), one series shifted forward relative to the other and with a small random component added. Their trajectory rotates counterclockwise (+per definition), Figure 1b. Since the LL method uses a moving window of three time steps, we can apply the method to time series that are not detrend and not stationary.

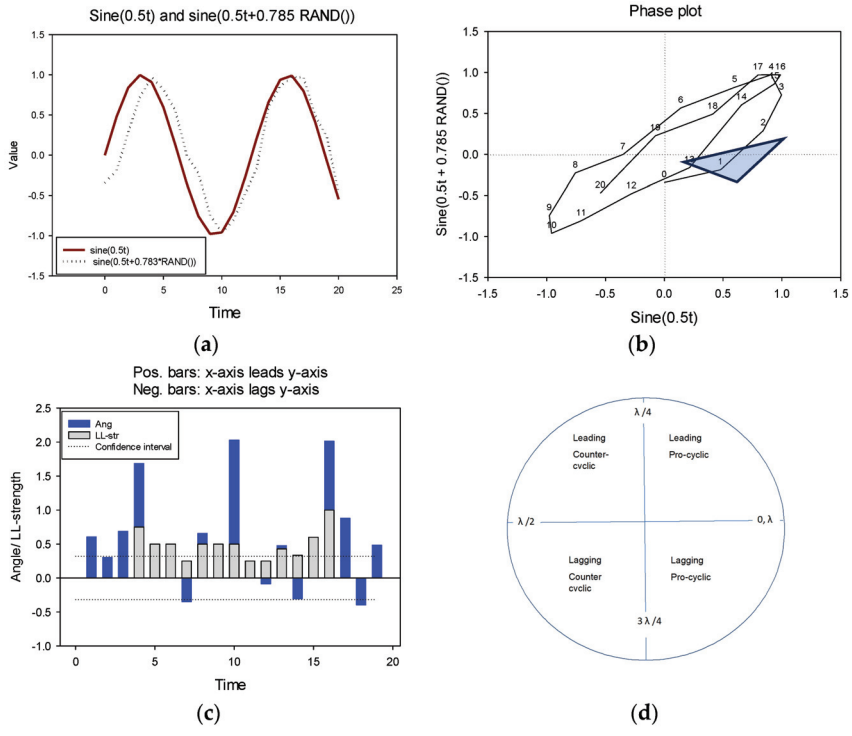


Figure 1. Example: Calculating lead-lagging (LL) relations and LL-strength. (**a**) Two sine functions: the smooth curve is a simple sine function, sin (0.5 t); the dashed curve has the form sin (0.5 t + ϕ × RAND()), where ϕ = +0.785. (**b**) In a phase plot with sin (0.5 t) on the x-axis and sin (0.5 t + ϕ RAND()) on the y-axis, the time series rotates; θ is the angle between two consecutive trajectories. The wedge suggests the angle between the origin and lines to observations 1 and 2. See text for details. (**c**) Angles between successive trajectories (black bars) and LL-strength (grey bars). Dashed lines suggest confidence limits for persistent rotation in the phase plot and persistent leading or lagging relations in the time series plot. Figure redrawn after (Seip et al. 2018). (**d**) LL relations and co-cyclic/counter-cyclic relations between two cyclic series as a function of the phase shift between them.

A measure of the persistence of a leading relation, the LL-strength, is defined as

$$LL\text{-}strength = (N_+ - N_-)/(N_+ + N_+)$$

where N_+ is the number of leading relations for two time series x and y, $x \rightarrow y$, and N_- is the number of lagging relations between the time series $x \leftarrow y$.

Thus, if there is a persistent positive LL relation over nine consecutive observations, the result is $N_+ = 9$, $N_- = 0$ and $LL = (9 - 0)/\ (9 + 0) = 1$. The number 9 is a tradeoff between measuring LL relations over short periods and the possibility of establishing significance. The angles, θ(3), are shown as blue bars in Figure 1c. The corresponding LL(9) values are shown as grey bars.

Confidence interval. The 95% confidence interval (CI) for the LL-strength is estimated using a Monte Carlo technique using two paired random time series. The confidence interval is found to be LL(9) = ± 0.32. Frequently, we will smooth time series, partly to avoid high-frequency noise and partly to identify cycle variabilities that are of interest, e.g., parliament election cycles. However, smoothing will increase the number of angles, θ(3), that show rotations in the same direction. Since we want to report the results for the smoothed versions, we use the term "pseudo-significant" when LL-strength <−0.32 or >0.32. Two cyclic series will either be leading or lagging in relation to each other. Visually, it is often seen as the peak of a leading series leading to a peak of the target series with less than half a common cycle period, $\lambda/2$, but the LL relation applies to all parts of the curve. If a leading series is inverted, it will be a lagging series. A leading series can both be pro-cyclic (b-coefficient positive) and counter-cyclic (b-coefficient negative) to the target series, Figure 1d.

Cycle periods, λ. When the trajectories complete one cycle in the phase plot, this corresponds to one cycle period in the time series plot. We calculate two measures of λ, one by calculating the running average λ over five consecutive time steps and one by taking the average λ of all cycle periods.

Power spectral density algorithm (PSD). The PSD algorithm identifies single cyclic series embedded in raw superimposed cyclic series. The cycle periods identified by the PSD method are compared to the cycle periods we identify with the LL method.

Lead and lag times or phase shifts (PS). The lead-lag shifts can be estimated by applying ordinary linear regression (OLR) to the time series. If they coincide perfectly, the regression coefficient, (r), is ≈ 1, and the lead-lag time, or PS, is zero. In Figure 1a, the two sine series are closely co-varying. Because both sine series were normalized to unit standard deviation, the b–coefficient (the slope) and "r" are both close to one. If the OLR gives $r = -1$, the two series are counter-cyclic, and the phase shift is around $\lambda/2$.

Principal component analysis (PCA). We apply PCA to five macroeconomic time series that describe the UK economy from 2004 to 2023. The PCA gives rise to two 2D plots: the score plot that shows how the economy changes over time and the loading plot that shows which of the five macroeconomic factors determine the position of the scores. For cyclic, non-Gaussian series, the position of variables in the loading plot will also reveal phase shifts between the series. The economic data were detrended to focus on events within a 3 to 5 years' time frame.

Inequality concerns. We use the terms "fairness", "GDP growth", "crime" and "war" to evaluate concerns for the effects of inequality.

We use the terms in Google Trends (score range 0 to ≈50), Google Scholar (score range 500,000 to ≈4 million) and Web of Science (score range 900 to ≈300,000). To evaluate and compare the strength of the concerns for each category, we center and normalize the scores to unit standard deviation for each source. However, other terms could also be used, e.g., "female empowerment", as in Lambert et al. (2003).

4. Results

The Google Trends time series for the 2004–2023 inequality data are shown in Figure 2a and, for conflict data, in Figure 2b. The graphs show the raw data as thin lines and

LOESS(0.4)-smoothed data as thick lines. Dropdown lines indicate the beginnings of the 2008 and the 2020 recessions.

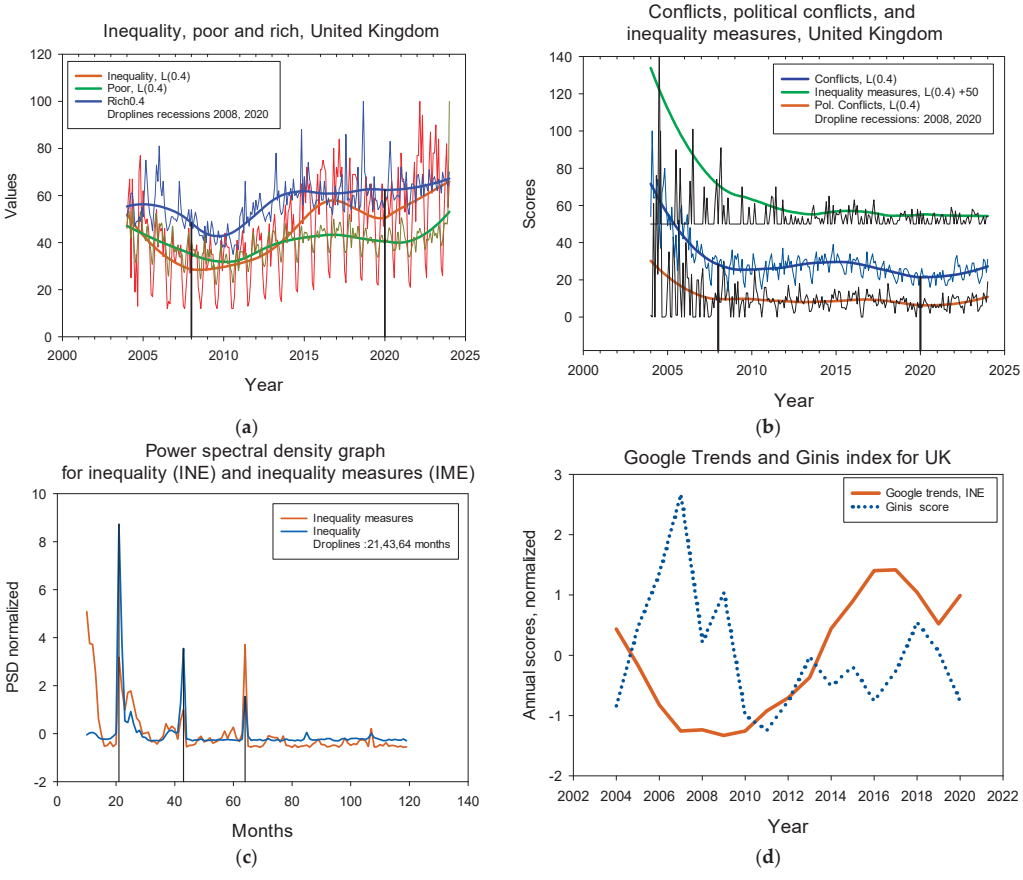


Figure 2. Google Trends. (**a**) Time series for inequality concerns. (**b**) Time series for conflicts and abatement measures. (**c**) Power spectral density graphs for inequality and inequality measures; common cycle periods are at 21 months, 43 months and 64 months. (**d**) Annual Google Trends for inequality, INE, compared to Gini's index for UK.

All three time series for inequality data—"inequality", "poor" and "rich"—showed similar patterns, and the decline in frequency during the two recessions was marked ("rich" may be an exception during the 2020 recession). All three time series for conflict data—"conflict", "political conflict" and "inequality measures"—levelled off until the 2008 recession. With the exception of "conflict", there were few non-zero data points in the series. However, both "conflict" and "inequality measures" showed an increase in frequency around 2015. We applied the PSD algorithm to the two time series, "inequality" and "inequality measures", and added drop lines for common cycle periods for the two series, Figure 2c. Common cycle periods for the two series occurred around 21, 43 and 64 months.

Based on these results, we LOESS-smoothed the two series so that they obtained cycle periods around 24 months (2 years) and 72 months (4 years), respectively.

Since the Gini index measures economic inequality in a country, we compared the Gini index for the UK with the Google Trends for INE for the period 2002 to 2020 (see Figure 2d). Since the two time series are measured in different units, we centered and

normalized the series. There do not appear to be any relations between the two series, p = 0.17. Chroufa and Chtourou (2022) suggest that the Gini index may not be the optimal measure of inequality in a GDP growth context.

The high-frequency movements. Figure 3a,b show the results for the high-frequency time series. The LL relations suggest that the two series are leading about equally often (43% pos., 56% neg.), Figure 3b. The cycle periods were around 25 months, and the PS were around 5 to 6 months (see Table 1).

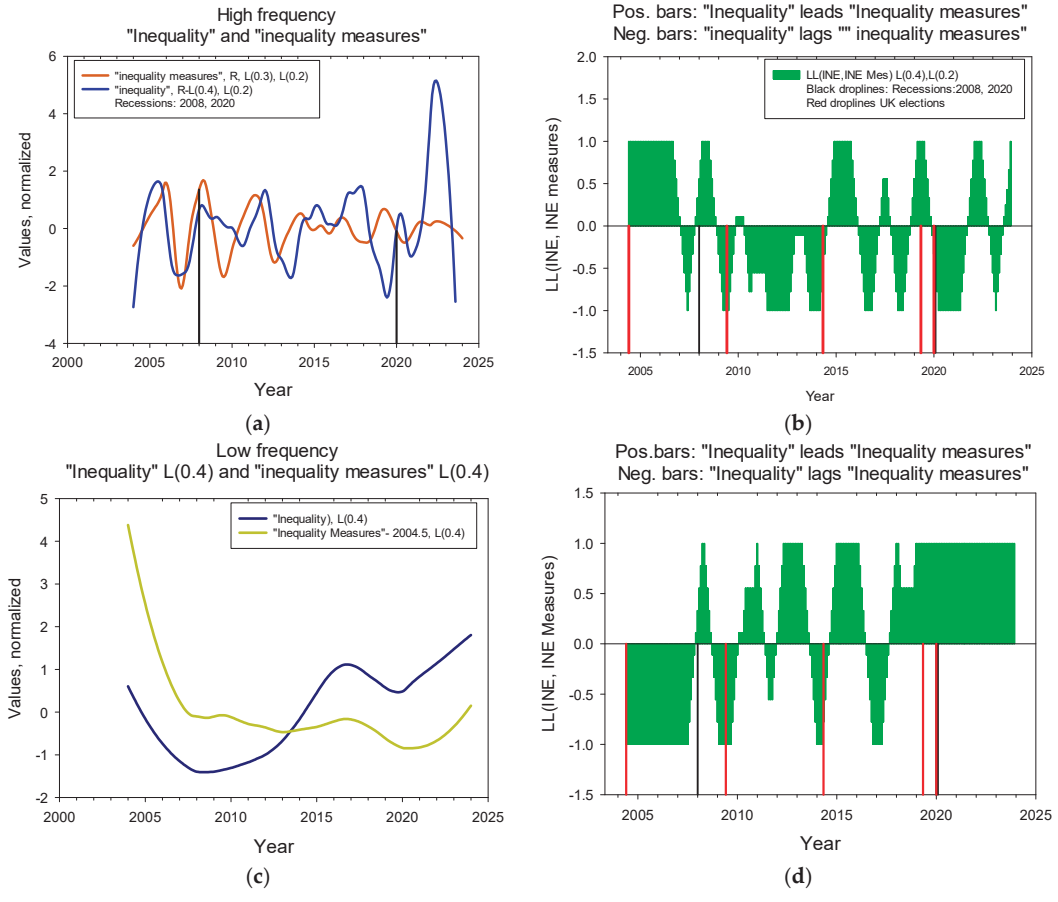


Figure 3. Cycle characteristics for detrended LL-relations between inequality and political conflicts. (**a**) Detrended and smoothed time series, high-frequency series. (**b**) Lead- lag relations between high-frequency series. The black droplines show the 2008 and the 2020 recessions; the red droplines show UK elections. (**c**) Detrended and smoothed time series, low-frequency series. (**d**) Lead-lag relations between low-frequency series.

Table 1. Common cycle period and phase shifts for high-frequency and low-frequency time series for the Google Trends records of "inequality" and "conflicts".

Variables	Cycle Periods (months)	Phase Shift (months)
HF	25	5–6
LF	190 ± 136	13–22

HF = high-frequency time series, LF = low-frequency time series.

The low-frequency movements. Figure 3c,d shows the LOESS(0.4)-smoothed series. The LL relations suggest that INE leads IME from about year 2010, Figure 3d. The cycle periods were, on average, 190 ± 136 months, and PS were, on average, from 13 to 22 months; see Table 1.

Figure 3a,b show, for example, LL relations for an "inequality" and an "inequality measures" series. The sequence of peaks and troughs in the series in Figure 3a can be compared visually to the calculated LL relations shown in Figure 3b.

4.1. "Inequality" and "Inequality Measures" during Recessions and Parliament Elections

The result of scores during recessions and elections to EU and UK parliaments are shown in Figure 4. For each month, a value greater than half the average frequency, denoted INE+, IME+ and LL+, would give a score of "1"; otherwise, it gives a score of "0".

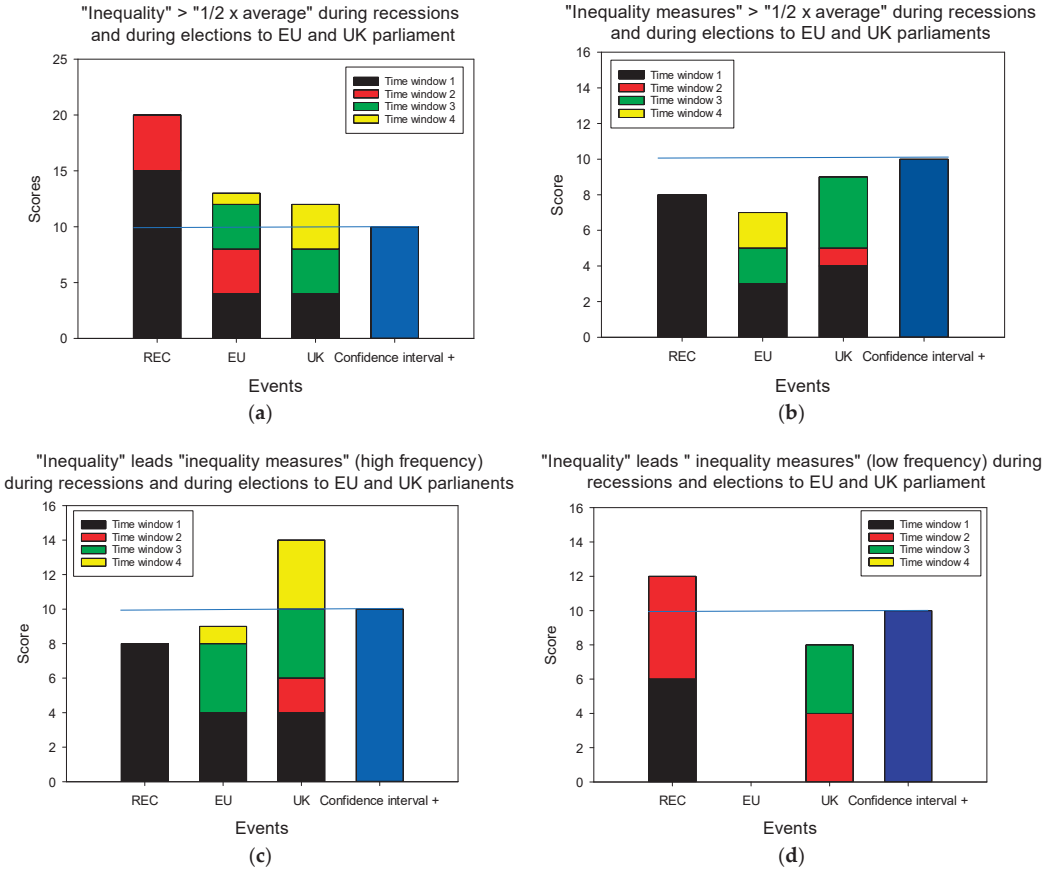


Figure 4. Number of months where "inequality" (INE) and "inequality measures" (IME) show positive scores during recessions and during elections to EU or to UK parliaments. (**a**) INE shows positive scores. The blue bar shows upper confidence interval for the three left columns and the blue horizontal line suggests confidence limits for the three left columns. (**b**) IME show positive scores. The blue bar shows upper confidence interval for the three left columns (**c**) INE leads IME, and LL(INE, IME) show positive scores for high-frequency series. The blue bar shows upper confidence interval for the three left columns (**d**) NE leads IME, and LL(INE,IME) show positive scores for high-frequency series. The blue bar shows upper confidence interval for the three left columns. The confidence interval for the EU and UK election events are between 2.23 and 2.39.

There are two time windows for recessions: one starting in 2008 and lasting for 16 months and one starting in 2020 and lasting for 6 months. There are 4 months for each time window for elections to EU and UK parliaments. In the columns, the time windows are assigned different colors.

A random distribution of scores would give about two for each of the election time windows. The blue column suggests a rough estimate of the upper confidence limit for the two recessions and four elections. Figure 4 shows the scores during three "events".

For the overall scores, the term "inequality" scores above the confidence limit during the three events: recessions, and elections to both EU and UK parliaments. IME scores below the confidence limit during all events. High-frequency INE leads IME during the UK elections, Figure 4c, and low-frequency INE leads IME during recessions, Figure 4d. It is the recessions and the UK elections that elicit the most pronounced response to inequality issues.

4.2. The UK Economy

The five time series that describe the UK economy are shown in Figure 5a.

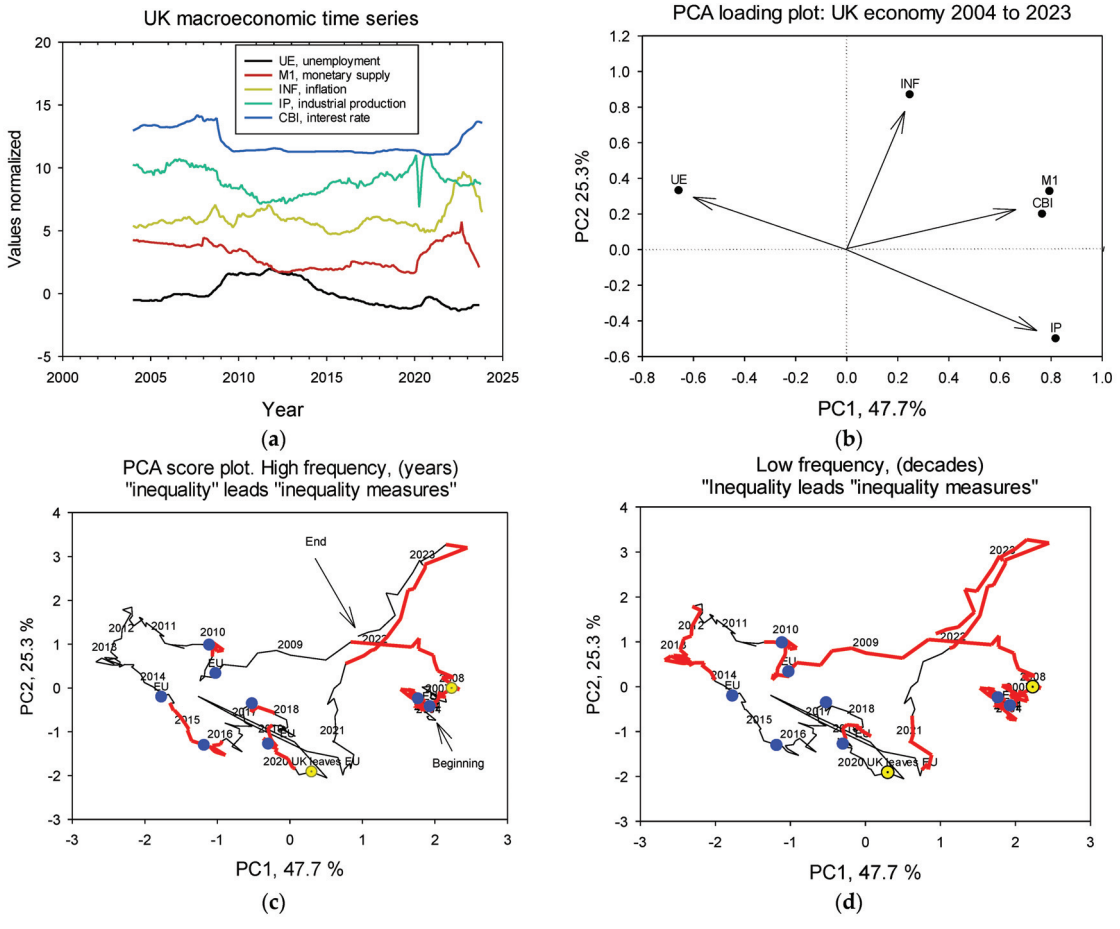


Figure 5. *Cont.*

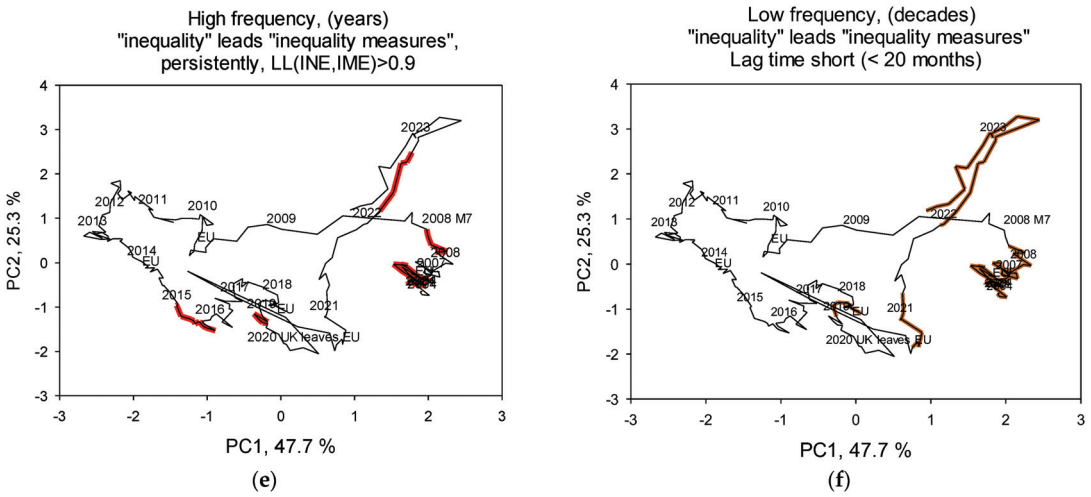


Figure 5. United Kingdom economy and elections to EU and UK parliaments. (**a**) Macroeconomic time series centered and normalized to unit standard deviation. (**b**) PCA loading plot for UK economy described by the five time series in (**a**). (**c**) PCA score plot for UK economy. Red lines show when "inequality" (INE) leads "inequality measures" (IME) with high-frequency time series. (**d**) PCA score plot for UK economy. Red lines show where INE leads IME with low-frequency time series. (**e**) Same as (**c**), but the leading relation is strongly established, LL(INE,IME) >0.9 (range 0–1.0), (**f**) Same as (**d**), but the lead time between INE and IME is short. (range 0–100 months). UE = unemployment, M1 = monetary supply, INF = inflation expressed as consumer price index, CBI = central banks short-term interest rate. Blue dots are elections in UK to the UK or EU parliaments. Yellow dots show the beginnings of recessions 2008 and 2020.

The PCA plots for the UK economy are shown as a loading plot in Figure 5b and as score plots in Figure 5c–f. The PCA score and loading plots are interpreted as follows: The position of an observation in a score plot (in our case, the state of the UK economy at a certain year) is associated with the position of the explanatory variables in the loading plot (in our case, the values of the five macroeconomic variables). Blue dots show parliament elections and yellow dots show recessions. The loading plot shows that the time series for the five time series matches a traditional pattern for an economy.

For example, unemployment (UE) varies inversely to industrial production (IP), as it should according to Okun's law, e.g., Maza (2022), Seip and Zhang (2022).

High-frequency results. The red lines in the left score plots in Figure 5c,e show the time windows where the high-frequency INE is a leading variable to the high-frequency IME.

The graph in Figure 5d shows results when LL relations are positive (>0; range −1 to +1), and the graph in Figure 5c shows results when LL relations are greater than 0.9, that is, the requirement for persistently tighter relations between INE and IME is strengthened.

Visually, the plot suggests that INE leads IME during periods just before, during or just after an election event, and the tighter requirement emphasizes the recession and elections during the years 2004 to 2008.

Low-frequency results. Figure 5d,f show scores where the low-frequency INE leads the low-frequency IME. The graph in Figure 5d shows the result when LL relations are positive, but for the graph in f, we have added the restriction that the lead time between INE and IME should be less than 20 months (range 0 to 100 months). The red curves in the upper right part of Figure 5d corresponds to states where INF, CBI and M1 have high values. (M1 is a technical variable expressing the amount of money in the society.)

High values of M1 and CBI suggest that the Bank of England tries to slow down the economy or reduce INF. The red curve in the upper left part corresponds to economic states where UE is high.

The graph in Figure 5f with the tighter requirement between INE and IME shows that the years from 2004 to 2008 are emphasized together with the time after the UK left the EU and the inflation period after the pandemic. The period 2011–2014 with high UE is no longer marked as an important event for discussing inequality issues. Figure 5c–f show that inequality concerns were less pronounced during the actual pandemic in 2020 than during the following inflation period up to 2023.

4.3. Inequality Concerns

Figure 6 shows two graphs that identify concerns that can be related to inequality. The first graph is based on the worldwide use of the concern terms—"fairness", "GDP growth", "crime" and "war"—in Google Trends, Google Scholar and the Web of Science. The second graph shows a PCA plot for the five Google Trends series: *inequality* and the four concern terms in the UK.

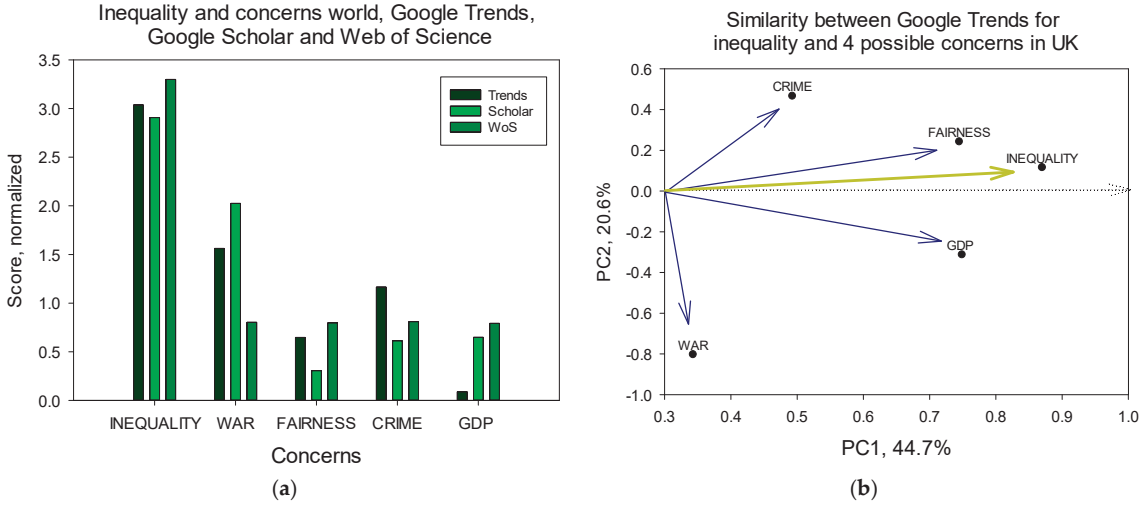


Figure 6. Concerns for the effects of inequality. (**a**) Concerns worldwide. "Trends" is Google Trends; "Scholar" is number of results in Google Scholar; "WoS" is number of results in Web of Science. (**b**) Google Trends 2004 to 2022. Co-movement between "inequality" and "crime", "fairness", "GDP growth" and "war".

5. Discussion

We first outline the general economic history of the UK. We thereafter discuss when the frequency of inequality concerns is high. Third, we discuss the results that show when discussions on inequality are followed by discussions on inequality abatement measures, expressed by a positive lead-lag relation LL(INE, IME) value. Fourth, we examine the possible consequences of inequality, such as sentiments of unfairness in the UK population and increases in crime. Last, we discuss the use of Google Trends and the HRLL methodology.

5.1. UK Economic History 2008–2023

A rough, stylized sketch of the UK economic history can be read by comparing trajectories in the PCA "map", Figure 5c–f, with the position of the macroeconomic variables in Figure 5b.

As trajectories in the panels of Figure 5c–f move towards the position for an acronym in panel Figure 5b, the variables represented by that acronym increase in importance for the UK economy.

Following the abrupt beginning of the recession in 2008, the economy, IP, slowed down and UE increased until 2012–2013, then UE fell, and IP grew until the COVID-19 pandemic recession in 2020. Following that recession (and the UK leaving the EU), INF rose rapidly until 2022 and then slowed down in about the same phase until the present.

It is interesting that combating the 2008 and the 2020 recessions resulted in quite different trajectories. The trajectory following the 2008 recession resulted in high UE, whereas the trajectories following the 2020 recession and pandemic resulted in high INF. However, the two recessions were handled differently using different monetary and fiscal policies.

5.2. High-Intensity Discussions of Inequality Issues

Interest in inequality, expressed as the relative frequency of the search terms "inequality", "poor" and "rich" as reported by Google Trends, were shown in Figure 2. The term "INE" declines before the recession but increases after the recession in 2008 for 6 to 9 months. The reason may be that there is an increase in employment before a recession but a "jobless recovery" after the recession (Seip and Zhang 2022). INE is on preponderance of probability or better associated with elections to the EU and UK parliaments over short-term time spans $\approx$ 24 months. "Preponderance of probability" here means that "inequality" and "inequality measures" are mentioned with higher frequency than the average during all 4 months around an election. The term IME was only associated with UK elections. Our hypothesis, H1, was therefore only partly supported; both INE and IME were used more frequently than expected (random) during EU and UK elections. However, IME was used less than INE, and the terms were not used persistently during all eight election events. Inequality and its abatement measures were discussed most consistently during the EU elections in October 2004 and the UK elections in May 2005 (black EU and UK columns in Figure 4).

5.3. Lead-Lag Relations

The term INE was leading IME during elections both to the EU and the UK parliament over short time spans. By embedding the results in the "map" of the UK economy, it is seen that INE leads IME both during elections and during the recessions in 2008 and in 2020. Over the low-frequency time spans, $\approx$72 months, the economic "map" showed that INE was leading IME during periods with high unemployment and high inflation and during recessions. Requiring that the LL relations between INE and IME are tight, the intensity of inequality conversations around the 2008 recession and the inflation period after the COVID-19 pandemic were emphasized. Two reasons may be important: (i) there is an increased urgency to achieve inequality abatement measures, or (ii) the driving force must last for some time (>20 months), corresponding to the cutoff value for the time between high INE and IME frequencies. Thus, our hypotheses H2 was supported—a leading role for INE to IME is related to political events, that is, either to elections for the UK or the EU parliaments or to harmful economic conditions expressed by unemployment and inflation. For all terms, we tried several synonyms, but all alternatives failed to produce sufficient data.

5.4. Inequality Concerns and Inequality Abatement Measures

Inequality concerns address the effects of inequality that reduce the quality of life or the affluence of the society. Abatement measures address what policies can be enacted to reduce either "objective" inequality, (e.g., expressed by the Gini index) or "subjective" inequality. Lambert et al. (2003, p. 1073) and Davidescu et al. (2024) include macroeconomic variables (GDP per capita) but also socioeconomic variables, such as public expenditures (schools) and gender policies, to explain subjective and real inequality concerns.

5.4.1. Inequality Concerns

There are two dichotomies with respect to "inequality". One set is if the inequality concerns relate to socioeconomic stress among the lower income individuals or if it relates to a discussion of measures to convince the rich to contribute more to the welfare system. The other set is if the rich get richer by unfair means and luck or by merit. We believe that concerns for socioeconomic stress would be dominant during recessions and during inflation events. Furthermore, discussions on stress would probably be more frequent on the internet, and discussion on how the rich could contribute more to welfare economy would probably be more frequent in reports and government hearing notes. However, in the UK, one could anticipate that discussion about contributions from the rich could be frequent during the Prime Minister period of the strongly conservative Liz Tuss from 6 September to 25 October 2022, (Tosun and Lucey 2023), but there was no pattern in the Google Trends that distinguished the period. To examine what type of concerns are most associated with inequality, we tried to combine the term "inequality" with the concern terms "fairness", "GDP growth", "crime" and "war", but UK Google Trends just reported "not sufficient data". However, time series were successfully established for the world. Figure 6a showed a comparison of the average frequency of the concern terms in Google Trends, Google Scholar and Web of Science. The first relates to public conversations, the second to scholarly research and discourses and the third to scientific publications. The concern most associated with inequality worldwide was *war* and, thereafter, *crime*. This holds for Google Trends, as well as for Google Scholar. For the Web of Science, all four concerns yielded similar results (the bars are of equal height).

For the concern terms in the UK (not paired with inequality), we obtained full series for all terms. Comparing the time series for the four terms to the series for inequality with a PCA analysis, we found that the series for fairness and GDP growth were most like the time series for inequality, Figure 6b. The overall results indicate that during elections, inequality gives reasons for discussions of abatement measures, and *fairness* and *GDP growth* are the main concerns during the discussions.

For the inequality concerns, like fairness, GDP growth, crime and war, most of the available literature gave information that was generic. The effects of inequality on economic growth were studied by Naguib (2015, p. 38 Appendix). The author examined countries that are members of the Organization for Security and Co-operation in Europe (OSCE) and found a statistically significant positive effect from inequality on GDP growth (a 1% increase in inequality gave a 1.2–1.5% increase in GDP, and the UK is included in the sample). In a model study, Lambert et al. (2003, pp. 1078, 1079) found that inequality aversion increased with the growth rate until it reached about 2%, but that growth above 2% reduced inequality aversion. It is not clear why it would not decrease persistently with growth rate. In economic terms, the Gini index for optimal growth rate is 38.2%; Lambert et al. (2003, Figure 2; model study) found it to be higher than the UK. Gini is 35.3 ± 1.2 for the period 2004 to 2020. Kelly (2000, p. 533) found that inequality had a strong and robust impact on violent crime in the United States, and Nafziger and Auvinen (2002) summarized findings, including their own studies, and found that inequality exacerbates the vulnerability of populations to humanitarian emergencies (war). However, the study addressed a selection of developing countries and thus did not include countries such as the UK.

5.4.2. Inequality Abatement Measures

Inequality abatement measures are outside the scope of the present study. However, during the COVID-19 pandemic in 2020, many countries implemented government support measures that alleviated the economic effects of the pandemic, (e.g., in Sweden, Angelov and Waldenström (2023)), and this may explain why the focus on inequality was less during the actual pandemic and stronger during the following inflation period. An interesting concept of a "natural rate of inequality" has been put forward by Lambert et al. (2003). The natural rate may refer to inequality sentiments of a population or to an optimal output rate.

5.5. How Economic Policies May Create Greater Inequality

Among economic states that solicit strong discussions of inequality are inflation and unemployment. Among the effects listed in the literature that would increase inequality are (i) increased profits for firms, (ii) reduced trade among countries and (iii) increased innovations. A micro-mechanism that could cause increasing inequality is the differences in consumer baskets for low- and high-income people (US data, Jaravel (2021, pp. 603, 605)) with households headed by single woman at the low end (OECD data, Azzollini et al. (2023)). Food is a larger part of the basket (food and energy prices tend to increase more than the average in a consumer basket). Finally, Filippin and Nunziata (2019, p. 119) suggest that perceived inflation is higher than actual inflation, but that there is a "keeping up with the Jones" effect along the whole income distribution.

To reduce inflation, a key tool for the central banks is to increase their short-term interest rates. However, increasing the interest rate may be associated with a *higher profit* for large firms and, again, affect low-income people more than high-income people through the consumer basket argument (Weber and Wasner 2023). *Increasing trade* may have contrasting effects on inequality. Low prices on traded goods, e.g., tools and machinery from China, will in principle favor low-income people, but the results do not seem to support this conjecture; see, for instance, Jaravel (2021, pp. 600, 6011, 6015). However, Rajaguru et al. (2023, p. 487) suggest that economic globalization aggravates income inequality (and led to the Brexit vote in 2016). Barth et al. (2023, p. 11) examined the political parties' election platforms for 169 European countries and found that increased import exposure decreased the welfare state support.

Innovations and high patent frequency may increase the demand for skilled (and educated) workers, whereas the demand for unskilled workers decreases (Díaz et al. 2020); (Jaravel 2021, p. 600). Since the unskilled workers belong to the low-income group, inequality would increase.

5.6. The Method

Most studiers of inequality address inequality and its effects by comparing effects of inequality among several countries; thus, ordinary linear regression (OLR), multiple regressions (MR) and panel data techniques are used. For example, Lambert et al. (2003) examined inequality across 96 countries, and Malla and Pathranarakul (2022) examined inequality across 68 countries. Some studies strengthen causality interferences by applying the Granger causality, (Granger 1969) or cross-correlation techniques (Kestin et al. 1998) to their data sets. However, both techniques require long data sets ($\approx$30 samples) and thus often find bi-directional causalities, e.g., the Ogbeide and Agu (2015) study on poverty and inequality in Nigeria and Cetin et al. (2021) on income inequality and technological innovation. If we averaged over long time series, we would also have found bi-directional causality for our time series; see Figure 3.

5.7. Robustness

Our focus was on the terms inequality and inequality abatement measures, but we could have searched for additional terms describing the effects of inequality as an issue in the political conversation. However, we found no terms that better described our intention with the study and that gave significant Google Trend series. We used the terms "fairness", "GDP growth", "crime" and "war" to identify the concerns associated with inequality. The terms were selected by comparing them to other similar terms in the Microsoft Word thesaurus. We originally wanted to use the term "morale", but fairness gave a more complete time series.

The HRLL method we use has been applied to sine functions with equal cycle periods but shifted in time relative to each other. It is then easily seen that it works as intended. However, in an application to forecasting algorithms in economics, it identified the forecast series as leading the observation about 80% of the time, and the economy was shown to be

anomalous when the forecasting series was not leading (Seip et al. 2019). Thus, we believe that the HRLL method identifies real (observed) LL relations correctly.

5.8. Further Work

The terms we use are exploratory, and it may be possible to find terms that better express people's sentiments. Our results for the UK could be generalized to other countries. For example, we downloaded inequality expressions for the US and found that several terms that did not deliver Google time series for the UK gave adequate time series for the US. Further studies should address abatement measures for unwanted consequences of inequality. A third issue is if it is possible to replace the model study by Lambert et al. (2003) on a "natural rate of inequality" by an empirical investigation based on UK data. Finally, we have discussed inequality on interannual and decadal scales, but inequality increases in many countries over multidecadal scales, and this could be the objective of further studies.

6. Conclusions

Inequality among people is a challenging issue in many countries and is hypothesized to cause political conflicts around themes including fairness, economic growth, crime, and war. In contrast to most other studies on inequality, we study the timing and strength of interest in a single country, the United Kingdom. We show, using Google Trends 2004 to 2022, that the term "inequality" precedes the term "inequality measures" around UK and EU parliament elections and during periods with unfavorable economic conditions (e.g., high inflation). Our results suggest that abating unwanted effects of inequality would be effective around parliament election times and when inflation and unemployment is high. However, since inflation and unemployment are the results of economic policy choices, it may be possible to implement abatement measures before inequality issues become serious.

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

Funding: This research was funded by Oslo Metropolitan University, grant number 1.

Informed Consent Statement: Not applicable.

Data Availability Statement: All data and all calculations are available from the first author.

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

References

Angelov, Nikolay, and Daniel Waldenström. 2023. COVID-19 and income inequality: Evidence from monthly population registers. *The Journal of Economic Inequality* 21: 351–79. [CrossRef]

Azzollini, Leo, Richard Breen, and Brian Nolan. 2023. Demographic behaviour and earnings inequality across OECD countries. *The Journal of Economic Inequality* 21: 441–61. [CrossRef]

Barth, Erling, Henning Finseraas, Anders Kjelsrud, and Kalle Moene. 2023. Openness and the welfare state: Risk and income effects in protection without protectionism. *European Journal of Political Economy* 79: 102405. [CrossRef]

Berisha, Edmond, David Gabauer, Rangan Gupta, and Chi Keung Marco Lau. 2021. Time-varying influence of household debt on inequality in United Kingdom. *Empirical Economics* 61: 1917–33. [CrossRef]

Bratanova, Boyka, Steve Loughnan, Olivier Klein, and Robert Wood. 2016. The rich get richer, the poor get even: Perceived socioeconomic position influences micro-social distributions of wealth. *Scandinavian Journal of Psychology* 57: 243–49. [CrossRef] [PubMed]

Burns, G. W., and W. C. Mitchell. 1946. *Measuring Business Cycles*. Cambridge, MA: National Bureau of Economic Research.

Cetin, Murat, Harun Demir, and Selin Saygin. 2021. Financial Development, Technological Innovation and Income Inequality: Time Series Evidence from Turkey. *Social Indicators Research* 156: 47–69. [CrossRef]

Chroufa, Mohamed Ali, and Nouri Chtourou. 2022. Inequality and Growth in Tunisia: New Evidence from Threshold Regression. *Social Indicators Research* 163: 901–24. [CrossRef]
Connor, Paul, Vasilis Sarafidis, Michael J. Zyphur, Dacher Keltner, and Serena Chen. 2019. Income Inequality and White-on-Black Racial Bias in the United States: Evidence from Project Implicit and Google Trends. *Psychological Science* 30: 205–22. [CrossRef] [PubMed]
Davidescu, Adriana AnaMaria, Tamara Maria Nae, and Margareta-Stela Florescu. 2024. From Policy to Impact: Advancing Economic Development and Tackling Social Inequities in Central and Eastern Europe. *Economies* 12: 28. [CrossRef]
Díaz, Guillermo Arenas, Andrés Barge-Gil, and Joost Heijs. 2020. The effect of innovation on skilled and unskilled workers during bad times. *Structural Change and Economic Dynamics* 52: 141–58. [CrossRef]
Filippin, Antonio, and Luca Nunziata. 2019. Monetary effects of inequality: Lessons from the euro experiment. *The Journal of Economic Inequality* 17: 99–124. [CrossRef]
Fisman, Raymond, Pamela Jakiela, and Shachar Kariv. 2017. Distributional preferences and political behavior. *Journal of Public Economics* 155: 1–10. [CrossRef]
Gong, Yuhan, Tim Li, and Lin Chen. 2020. Interdecadal modulation of ENSO amplitude by the Atlantic multi-decadal oscillation (AMO). *Climate Dynamics* 55: 2689–702. [CrossRef]
Granger, C. W. J. 1969. Investigating Causal Relations by Econometric Models and Cross-spectral Methods. *Econometrica* 37: 424. [CrossRef]
Hoy, Christopher, Russell Toth, and Nurina Merdikawati. 2024. A false divide? Providing information about inequality aligns preferences for redistribution between right- and left-wing voters. *The Journal of Economic Inequality* 2024: 1–39. [CrossRef]
Jaravel, Xavier. 2021. Inflation Inequality: Measurement, Causes, and Policy Implications. *Annual Review of Economics* 13: 599–629. [CrossRef]
Kelly, Morgan. 2000. Inequality and Crime. *The Review of Economics and Statistics* 82: 530–39. [CrossRef]
Kestin, Tahl S., David J. Karoly, Jun-Ichi Yano, and Nicola A. Rayner. 1998. Time–Frequency Variability of ENSO and Stochastic Simulations. *Journal of Climate* 11: 2258–72. [CrossRef]
Lambert, Peter J., Daniel L. Millimet, and Daniel Slottje. 2003. Inequality aversion and the natural rate of subjective inequality. *Journal of Public Economics* 87: 1061–90. [CrossRef]
Malla, Manwar Hossein, and Pairote Pathranarakul. 2022. Fiscal Policy and Income Inequality: The Critical Role of Institutional Capacity. *Economies* 10: 115. [CrossRef]
Maza, Adolfo. 2022. Regional Differences in Okun's Law and Explanatory Factors: Some Insights from Europe. *International Regional Science Review* 45: 555–80. [CrossRef]
Michalek, Anton, and Jan Vybostok. 2019. Economic Growth, Inequality and Poverty in the EU. *Social Indicators Research* 141: 611–30. [CrossRef]
Nafziger, E. Wayne, and Juha Auvinen. 2002. Economic Development, Inequality, War, and State Violence. *World Development* 30: 153–63. [CrossRef]
Naguib, Costanza. 2015. The relationship between inequality and GDP growth: An empirical approach. In *LIS Working Paper Series*. Luxembourg: Luxembourg Income Study (LIS).
Ogbeide, Evelyn Nwamaka Osaretin, and David Onyinyechi Agu. 2015. Poverty and Income Inequality in Nigeria: Any Causality? *Asian Economic and Financial Review* 5: 439–52. [CrossRef]
Rajaguru, Gulasekaran, Sadhana Srivastava, Rahul Sen, and Pundarik Mukhopadhaya. 2023. Does globalization drive long-run inequality within OECD countries? A guide to policy making. *Journal of Policy Modeling* 45: 469–93. [CrossRef]
Roe, Mark J., and Jordan I. Siegel. 2011. Political instability: Effects on financial development, roots in the severity of economic inequality. *Journal of Comparative Economics* 39: 279–309. [CrossRef]
Rözer, Jesper, Bram Lancee, and Beate Volker. 2022. Keeping Up or Giving Up? Income Inequality and Materialism in Europe and the United States. *Social Indicators Research* 159: 647–66. [CrossRef]
Sandnes, Frode Eika, Anders Örtenblad, and Einar Duengen Bøhn. 2023. For or Against Equal Pay? A Study of Common Perceptions. *Compensation & Benefits Review* 55: 87–110. [CrossRef]
Seip, Knut L., Oyvind Gron, and Hui Wang. 2018. Carbon dioxide precedes temperature change during short-term pauses in multi-millennial palaeoclimate records. *Palaeogeography, Palaeoclimatology, Palaeoecology* 506: 101–11. [CrossRef]
Seip, Knut Lehre, and Dan Zhang. 2022. A High-Resolution Lead-Lag Analysis of US GDP, Employment, and Unemployment 1977–2021: Okun's Law and the Puzzle of Jobless Recovery. *Economies* 10: 260, Corrected in *Economies* 11: 75. [CrossRef]
Seip, Knut Lehre, and Robert McNown. 2007. The timing and accuracy of leading and lagging business cycle indicators: A new approach. *International Journal of Forecasting* 23: 277–87. [CrossRef]
Seip, Knut Lehre, Yunus Yilmaz, and Michael Schröder. 2019. Comparing Sentiment- and Behavioral-Based Leading Indexes for Industrial Production: When Does Each Fail? *Economies* 7: 104. [CrossRef]
Timoneda, Joan C., and Erik Wibbels. 2022. Spikes and Variance: Using Google Trends to Detect and Forecast Protests. *Political Analysis* 30: 1–18. [CrossRef]
Tosun, Onur Kemal, and Brian Lucey. 2023. Growth... What growth? *Finance Research Letters* 52: 103594. [CrossRef]

Weber, Isabella M., and Evan Wasner. 2023. Sellers' inflation, profits and conflict: Why can large firms hike prices in an emergency? *Review of Keynesian Economics* 11: 183–213. [CrossRef]
Wilkinson, Richard, and Kate Pickett. 2010. *The Spirit Level: Why Greater Equality Makes Societies Stronger*. New York: Bloomsbury Press.

economies

Article

Patterns of Intergenerational Educational (Im)Mobility

Enzo Valentini

Department of Political Science, Communication and International Relations, University of Macerata, 62100 Macerata, Italy; enzo.valentini@unimc.it

Abstract: Intergenerational education mobility is a key dimension of social mobility and explores the extent to which educational attainment is transmitted across generations within a society. The implications of low education mobility concern both equity (everyone should have the same opportunities) and efficiency (it would be good for the economy and society if the most gifted and deserving young people were to study and not the children of the already educated). The literature identifies several drivers that can influence the level of social mobility in general and education mobility specifically, including characteristics of educational systems, public spending, degree of urbanisation, informal frictions, and beliefs. This paper seeks to identify 'patterns of intergenerational education (im)mobility' through a cluster analysis that takes into account the level of intergenerational mobility in education and a number of variables concerning its possible drivers, considering data on 82 countries (with different levels of development). The advantage of cluster analysis lies in the possibility of identifying regularities, but avoiding reasoning 'on average', i.e., safeguarding the possibility that different social patterns may exist. The results also allow us to speculate on possible policies to increase school mobility, highlighting, among other things, the 'equalising' role played by public spending on education.

Keywords: intergenerational education mobility; social mobility; cluster analysis; public spending in education

Citation: Valentini, Enzo. 2024. Patterns of Intergenerational Educational (Im)Mobility. *Economies* 12: 126. https://doi.org/10.3390/economies12060126

Academic Editor: Gheorghe H. Popescu

Received: 27 March 2024
Revised: 13 May 2024
Accepted: 17 May 2024
Published: 21 May 2024

1. Introduction

The study of the intergenerational transmission of socio-economic statuses is interesting and topical in several respects.

On the one hand, it helps us delve into the well-known theme of 'equality of what?' (Sen 1980). In fact, it is quite evident that a strong correlation of socio-economic statuses between parents and children is a sign of some lack of equality of opportunity. From this point of view, inequality of outcomes (e.g., income) can be more or less socially acceptable if it is accompanied, or not, by equality of opportunity. Nonetheless, high income inequality can itself be an obstacle to a real equality of opportunity (Corak 2013).

On the other hand, mobility across generations is not only an issue of equality, but also of efficiency, especially when defined in terms of educational mobility (D'Addio 2007). It is important that it is the most able and willing people who study (and gain access to the resulting social roles) and not (only) people from rich families. Otherwise, we would end up with people with inadequate abilities in key roles (indeed, as often seems to be the case), and we would lose the opportunity to socially take advantage of the abilities of gifted people just because they come from 'disadvantaged' families or social backgrounds (Glomm and Ravikumar 1992; Lloyd-Ellis 2000; Staffolani and Valentini 2007).

The focus of the research proposed in this paper is precisely on educational mobility, which among other things is considered to be a key element of overall economic mobility across generations (Feinstein et al. 2006; D'Addio 2007; Jerrim and Macmillan 2015; Narayan et al. 2018; Stuhler 2018).

The literature identifies several factors that may influence the level of educational intergenerational mobility (and, through it, intergenerational income mobility).

Public spending in education is a key element. It can not only improve the quality of education in general, but also benefit those who would be at risk of starting off disadvantaged. From a theoretical point of view, Solon (2004) proposes a model according to which intergenerational income elasticity increases as the return on investment in human capital increases, but which decreases with the progressivity of public investment in human capital. The model of Davies et al. (2005) analyses intergenerational earning mobility in a framework where human capital plays a crucial role and concludes that mobility is higher with public than with private education. Herrington (2015) developed an overlapping generations model and calibrated it in order to compare the US and Norway empirically and quantitatively, reaching the conclusion that public education spending plays a key role in intergenerational earning persistence. Lee and Seshadri (2019) present a model of human capital investments that explains the intergenerational persistence of earnings, wealth, and college attainment and conclude that education subsidies can reduce the intergenerational persistence of economic status. Mayer and Lopoo (2008) focused on human capital investment. They used data from the Panel Study of Income Dynamics and the US Census of Governments and found greater intergenerational mobility in high-spending states compared with low-spending states. Neidhöfer et al. (2018) computed several indexes of intergenerational education mobility for 18 Latin American countries, finding significant cross-country differences also associated with public educational expenditures. Balcazar et al. (2015) used data on 48 countries that participated in the Programme for International Student Assessment (PISA) in 2012, finding evidence of a correlation between public spending on schooling and inequality of opportunity in achieving basic proficiency in reading, mathematics, and science. In Narayan et al. (2018), regressions involving data on richer economies showed that higher public spending on education is associated with higher relative intergenerational mobility in education, and the authors concluded that '[t]his is consistent with the theory that public spending helps equalise opportunities through investments that compensate for the gap in private investments between children of rich and poor parents' (Narayan et al. 2018, p. 19).

Moreover, if public spending on education is specifically directed towards individuals from disadvantaged economic backgrounds and is financed through taxes that particularly fall on individuals from wealthy families (e.g., bequests taxation), it can be particularly effective in rebalancing opportunities for access to education itself, fostering equity and efficiency gains (Staffolani and Valentini 2007).

Public intervention in this field can also concern regulation, and the presence and duration (in years) of compulsory education can certainly influence educational mobility, particularly in the most disadvantaged contexts. An uneducated (and therefore more likely to be 'poor') parent might decide not to make their child study (or they may study less than they should). If a certain number of years of education is compulsory, this choice is limited. It is important to emphasise that the years of compulsory education in general are somewhat related to the level of public spending on education; however, they represent a slightly different aspect of public intervention directly related to regulation rather than the level of spending.

Socio-economic segregation is another factor that can strongly influence educational mobility. Van der Weide et al. (2021), using data on 153 countries, found that proxies for segregation are negatively correlated with intergenerational education mobility. A specific definition of this concept concerns spatial/residential segregation. Recent research suggests that more residentially segregated areas (i.e., where families with different socio-economic backgrounds and races live in separate neighbourhoods) tend to have lower intergenerational mobility. Chetty et al. (2014) used administrative records of more than 40 million children in the US and, with OLS (Ordinary Least Squares) regressions across 700 areas, found that high-mobility areas have less residential segregation. Connolly et al. (2019) linked microdata from the US and Canada, and their OLS estimates suggest that 'inequalities between whites and blacks likely play an important role in understanding why the United States has lower rates of intergenerational mobility' (Connolly et al. 2019,

p. 598). In that case, 'segregation' is to be understood in terms of 'social segregation' (which then often also becomes 'spatial' segregation). Corak (2021) used data from 266 Canadian Census Divisions on a cohort of men and women born between 1963 and 1970, finding that low-mobility regions tend to exist outside of urban areas; reading his results, he underlines that Durlauf and Seshadri (2018) explicitly model the influence that inequality has on economic opportunity through its impact on socio-economic segregation. In the present paper, residential segregation is mainly considered in terms of the degree of urbanisation, which seems especially relevant for developing countries; obviously, a lower degree of urbanisation indicates greater spatial segregation and social isolation.

As mentioned earlier, income inequality itself can trigger inequality of opportunity in access to education in a dangerous vicious circle. Corak (2013) hypothesises the existence of this risk with a theoretical approach that also refers to empirical results from the literature. Reardon (2011) carried out a descriptive study on the relationship between academic achievement and family income in the US over the last few decades using data from 12 nationally representative studies that included information on family income and student performance in math or reading. He found that socio-economic status is a predictor of student academic achievement and educational outcomes. Campbell et al. (2005) used microdata from the US census, and their estimates suggest that an increase in family income and wealth inequality leads to a growing dispersion of educational attainments. Kearney and Levine (2016) used US longitudinal microdata, and their regressions confirm that greater income inequality might lead to lower levels of high school completion among individuals from low-income families and, therefore, to lower rates of upward mobility. A study with US data showed that, with respect to income, about 50 percent of intergenerational persistence can be explained by parents' investments in their children's education (Restuccia and Urrutia 2004). This mechanism tends to favour access to education for individuals belonging to wealthy families. Duncan and Murnane (2012) found that, in the last few decades, the amount that high-income families spent on their children's education grew by 150 percent, while the amount spent by low-income families grew by 57 percent.

The relevance of cultural context and 'family culture' in influencing intergenerational mobility is often considered in the literature (Becker and Tomes 1979; Mayer 1997; Piketty 2000; see McLanahan 2020 for a very comprehensive review on the subject), but less so in the empirical literature. A factor not studied so far in the empirical literature lies in 'cultural attitudes' concerning the importance of children's independence. If parents consider it important for their children to be independent and self-sufficient, the link between parents' economic status and children's opportunities may be weakening. An extreme case that can be used to exemplify this concept is as follows: 'it is important to me that you are autonomous and independent, I do not pay for your studies, you have to work to pay for them yourself'.

The underlying idea of the analysis in this paper is that the factors listed so far (public spending on education, compulsory education, income inequality, segregation/urbanisation, and importance of children's independence) may be more or less important in different socio-economic contexts. These differences may especially emerge between advanced economies and developing countries (regarding the trends of intergenerational mobility over time in countries at different stages of development, see also Leone 2019).

Some of the literature suggests that mobility will decline as incomes increase in the absence of public interventions (Becker and Tomes 1979; Becker et al. 2018); therefore, the level of public spending on education could be more effective in fostering educational intergenerational mobility in advanced economies. The effect of income inequality on children's educational opportunities may be greater in rich countries, since the link may exist through the resources invested by parents in their children's education; the availability of these resources in addition to those needed to satisfy basic needs is certainly greater in rich economies. Compulsory education may be more important for developing countries because it has a greater influence on low and middle levels of education. The same applies to the degree of urbanisation because the difference between rural and urban environments

is the greatest cause of social segregation in developing countries. For example, Reddy and Singh (2021) found that, in India, the intergenerational persistence in educational attainment is much higher in rural areas than in urban areas.

It is conceivable that the cultural emphasis on child autonomy may have greater effects in advanced economies because there are actually more chances to be autonomous and independent (less unemployment, more educational institutions, more public services, more scholarships, etc.).

The average level of education (often related to per capita income) can also influence mobility and can do so differently in countries with different levels of development. 'In the world's poorest countries, a large majority of parents have no education. When parents are equally deprived, it matters less what household one is born into, implying a high level of relative mobility. As countries increase their education and income levels, the gaps between poor, middle-class and better-off parents become more pronounced. Without public interventions, children's education trajectories will eventually start to diverge depending on whether they are born into a poor, middle-class or upper-class family, which is when relative intergenerational mobility will decline' (Van der Weide et al. 2021, p. 30).

On this basis, this paper proposes a cluster analysis with data from 82 countries in an attempt to verify the existence of different patterns in the relationships between intergenerational mobility in education and the factors discussed above, with particular reference to differences between economies at different stages of development.

Most of the empirical analyses mentioned in this introduction, on the contrary, use individual data and a micro-econometric approach. Therefore, they are also able to deepen the analysis in terms of the causality of the identified relationships. This leads to the biggest limitation of the analysis proposed here: from a technical point of view, it does not allow us to make considerations regarding the causality of the relationships. We shall return to this matter in the subsequent parts of the article as well as in the closing remarks. The approach followed in this article, as already mentioned, uses aggregated data and carries out a more general, cross-country analysis with a cluster methodology.

Nevertheless, the approach followed in the present article has innovative and useful aspects that can improve our understanding of the analysed phenomenon. It allows us to consider together several factors that can influence intergenerational mobility in education, to also include among these factors the cultural aspect (independence of children) that is generally not considered in the empirical literature, and to carry out an analysis on a sample involving a very high number of nations, even at different levels of development, with a methodology that can highlight different relationships between these factors and mobility depending precisely on the level of development.

2. Methodology and Data

As highlighted in the introduction, the idea behind this paper is that the various factors possibly influencing intergenerational mobility in education may have different impacts in economies with different levels of development.

One way, though not the only way, in which this concept can be defined in an empirical analysis in which the units of analysis are national economies is as follows: is it possible to identify patterns (i.e., group countries) according to level of income, mobility in education, public spending on education, compulsory education, income inequality, level of urbanisation, and importance of children's independence?

Cluster analyses are one of the best ways to search for groups through data (Kaufmann and Rousseeuw 1990). In our case, it seems to be an appropriate way to cluster nations according to level of income and level of intergenerational mobility in education in the first instance and according to the other variables in the second instance. The number (82) of observations available in the dataset constructed for this analysis also suggests the choice of this methodology. If one thinks that the relationships between the variables may be different for groups of countries (e.g., high-, middle-, and low-income economies), the

number of available observations would not allow for reliable regressions on subgroups of the available sample.

This paper seeks to identify 'patterns of intergenerational education (im)mobility' through a cluster analysis. The advantage of cluster analyses lies in the possibility of identifying regularities but avoiding reasoning 'on average', i.e., safeguarding the possibility that different social patterns may exist.

Before exploring the details of the analysis technique, it is useful to describe the variables and data.

Table 1 shows the variables used in the analysis, the databases of origin, and, finally, the mean and standard deviations (for the 82 countries considered). All variables are averaged over a relatively long period of time. This aggregation over time smooths out the changes in and development of a country over 35 years. This may lead to an underestimation of any changes in trends during the period under analysis. On the other hand, it seems to be a good way to approximate the conditions in which people born in the 1980s grew up.

Table 1. Dataset description and summary statistics.

Variable	Source	Time	Variable (Shortened)	Mean	Standard Deviation
GDP per capita (Constant 2017 Int. Dollars)	World Bank World Development Indicators	(average value 1980–2015)	GDP	USD 20197	15,392
Educational attainment (% of the population 25+ who at least completed upper secondary school)	World Bank World Development Indicators	(average value 1980–2015)	EDU	53.8%	24.0
Urban population (% of tot. population)	World Bank World Development Indicators	(average value 1980–2015)	URB	61.9%	18.1
Gini Index	World Bank World Development Indicators	(average value 1980–2015)	GINI	0.369	0.083
Government expenditure on education (% GDP)	World Bank World Development Indicators	(average value 1980–2015)	GOV_EDU	4.39%	1.36
Compulsory education (years)	World Bank World Development Indicators	(average value 1980–2015)	COMP_EDU	9.7	1.7
Share of parents indicating 'Children's Independence' as an important quality in children (%)	European Values Study—World Values Study Integrated survey 1981–2021	(average value 1981–2021) [1]	CHILD_IND	47.2%	14.3
Expected rank of a child whose parents rank in the bottom half of the education distribution	World Bank Global Database on Intergenerational Mobility [2]	1980s cohort (generation born between 1980 and 1989)	MOB_EDU	39.3	3.4

[1] Not all nations are involved in all EVS and WVS waves. Averages were calculated using the data available over the time period 1981–2021. For each available wave, the microdata were processed using the weights suggested by the relevant guidelines. [2] (GDIM 2023; Van der Weide et al. 2023).

All countries for which data are available are included in the analysis, with the sole exception of those classified by the World Bank (on 1 July 2020) as being in the 'Low Income' category in order to avoid data reliability problems (between 1980 and 2015, these nations had few values for some variables, so the values of those variables would not be true

'period averages'). These are almost exclusively countries from Sub-Saharan Africa (a region that is nevertheless represented in the dataset).

Table A1 in Appendix A lists the countries included in the analysis and the values of the variable concerning intergenerational mobility in education. Appendix B details the information required to retrieve data from the databases used to ensure the reproducibility of the analysis.

The variable related to intergenerational mobility in education is listed in the last row of Table 1 and deserves a closer look. It measures 'the expected rank of an individual in the education distribution whose parents rank in the bottom half of the parent education distribution' (Van der Weide et al. 2021, p. 8). The education of parents is identified by referring to the parent with the highest level of education, while, for the next generation, the entire population is considered (thus, both sons and daughters). This is a rank-based measure of relative mobility. Relative mobility represents the extent to which an individual's position in the distribution of educational attainments is independent of the position of his/her parents. If the educational attainment of each individual is independent of the starting conditions (i.e., parental status), the expected rank should be 50 for all individuals (both for those with parents who were in the bottom half of the ranking for their generation and for individuals with parents who were better placed in the same ranking as the previous generation). Hence, the variable represents the degree to which an individual's educational level is independent of the educational level of parents. A higher value of this variable is associated with greater intergenerational mobility (or, in other words, greater equality of opportunities).

Note that the low standard deviation of the 'compulsory education' variable results from 60 of the 82 countries having values between 9 and 12. This does not detract from the fact that there are countries with high values, such as The Netherlands, which has a value of 13 because full-time education is compulsory from the ages of 5 to 16 and at least part-time education is compulsory from the ages of 16 to 18. There are some very low values, such as Bangladesh (5), Malaysia (6), and Iraq (6). Note that these are not the latest values but the average of the values over the period of time considered calculated according to the availability of data, which often does not cover the entire period.

Cluster analyses are affected by the magnitude of the variables. It was therefore necessary to standardise (mean, 0; standard deviation, 1) the variables to make them comparable.

The intention was to group nations according to the values of these (standardised) variables, and this was achieved using a partition cluster analysis method (k-means), in which the observations (countries) are broken up into several nonoverlapping groups (Hastie et al. 2009). Each group brings together nations that share a common pattern in the mean values of the variables considered. The analysis used an iterative algorithm that, starting from k initial cluster centres, minimises the Euclidean distance within each cluster between its mean (k-clustering) and its observations while maximising the distance in terms of means among adjacent clusters. For this, we used the command 'cluster kmeans' in Stata, which implements an iterative procedure. It begins with k initial group centres, assigns observations to the group with the closest center, computes the mean of the observations assigned to each of the groups, and then repeats the process. These steps continue until all observations remain in the same group from the previous iteration.

The replicability of results deserves much attention in this type of analysis. Two important choices must be made: the optimal number of groups (k) must be chosen and the number of random starting points (r) must be determined. For both issues, the procedure suggested by (Makles 2012) was followed.

Given a certain number of random starting points, the optimal number of groups can be determined by comparing the values of the within sum of squares (WSS) or its logarithm for each k. In this case, one can stop at the number of groups k beyond which increasing the number of groups does not contribute significantly to the reduction in the WSS (i.e., fragmenting into further groups does not explain the data any better). It is also possible

to compare, for each *k*, the η2 coefficient, which is quite similar to the R2. One should stop at the number of groups k beyond which increasing the number of groups does not significantly increase the η2 coefficient. Finally, the proportional reduction in error (PRE) coefficient can be used. It illustrates 'the proportional reduction of the WSS for cluster solution k compared with the previous solution with k − 1 clusters' (Makles 2012, p. 347).

However, even the number of random starting points cannot be chosen arbitrarily. 'The best way to evaluate the chosen solution is therefore to repeat the clustering several times with different starting points and then compare the different solutions' (Makles 2012, p. 350).

Figure 1 reports the within sum of squares (WSS), its logarithm, the η^2 coefficient, and the proportional reduction in error (PRE) as k increases. Each line represents a different repetition of the analysis as the number of random starting points changes (from 100 to 150). Looking at the values of WSS and η^2, we see that the groups do not add much information when going beyond 4. However, when looking at the PRE, we see that, for up to k = 6, in quite a number of repetitions (i.e., varying the number of random starting points) the addition of another group contributes quite significantly to the reduction in the WSS. So, the most prudent choice seems to be k = 6. Finally, among all the repetitions with k = 6, the repetition selected had the lowest WSS and the greatest η^2 (in both cases, this occurred with 131 random starting points).

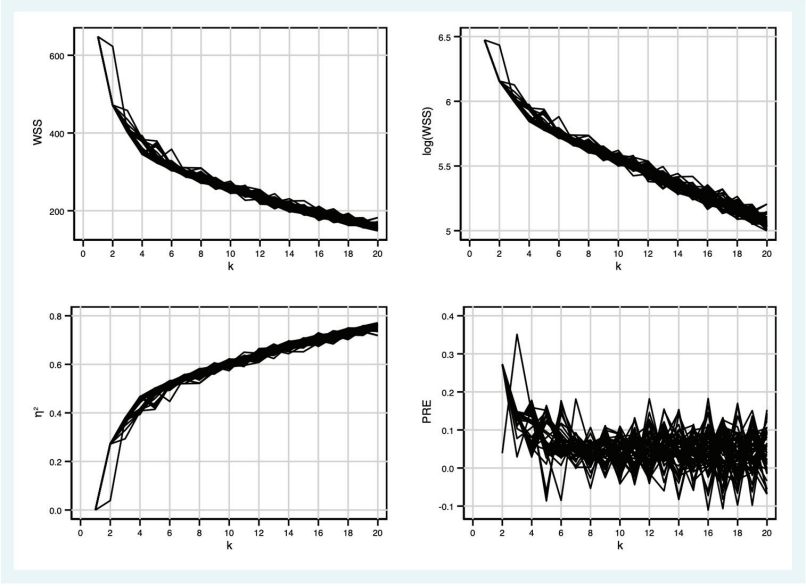


Figure 1. K-means cluster analysis. Selection of the optimal number of groups.

3. Results

The steps illustrated in the previous section suggest grouping the data into six groups with 131 random starting points via a k-means cluster algorithm. Table 2 shows the results of the resulting clusters/groups in terms of average values of the (standardised) variables.

To facilitate a possible reading of the outcomes, the clusters were ordered according to the level of GDP per capita. The idea is to interpret the results by comparing the variables' values between clusters with similar levels of development (High, Middle, and Low GDP). Cluster A must therefore be compared with cluster B, cluster C must be compared with cluster D, and cluster E must be compared with cluster F. The bold and underlined numbers indicate that the average value for that variable of the cluster is meaningfully higher than that of the cluster it is compared to.

Table 2. Cluster characteristics (mean values of standardised variables).

	Cluster	GDP	MOB_EDU	EDU	URB	GINI	GOV_EDU	COMP_EDU	CHILD_IND	N (Countries [1])
High GDP	A	1.31	0.19	0.58	0.74	**−0.50**	0.49	**0.86**	0.11	14
	B	1.23	**1.11**	**0.69**	0.77	−0.82	**1.03**	−0.05	**1.33**	11
Middle GDP	C	−0.37	**−0.04**	**0.85**	−0.07	−0.45	**−0.05**	−0.22	**−0.17**	22
	D	−0.54	−0.90	−0.76	**0.42**	**1.77**	−0.20	**0.83**	−0.80	12
Low GDP	E	−0.66	**1.50**	−1.13	**−0.80**	**1.05**	−1.05	**−0.52**	−1.13	5
	F	−0.77	−0.60	−1.10	−1.01	−0.04	**−0.52**	−0.77	**−0.10**	18

[1] Appendix C lists countries that fall under the different clusters. Bold and underlined numbers indicate that the average value for that variable of the cluster is meaningfully higher than that of the cluster in the same category of GDP per capita (High, Middle, and Low).

Cluster analyses are often used in the literature to generate hypotheses that should then be tested. In the present analysis, the procedure could be: (a) to group countries into clusters with different links between the variables (e.g., public spending on education is associated with higher intergenerational mobility in rich countries, whereas the relationship seems to be the reverse among poor countries); (b) to test the causality of the relationships (and its direction) in subgroups of countries based on the results of the cluster analysis (e.g., a causality analysis taking only nations from clusters A and B and another causality analysis taking only nations from clusters E and F). This would certainly be interesting, but it goes beyond the scope of this paper. Additionally, a purely technical issue lies in the number of nations in each subgroup, which is too small to carry out any analysis of causality (continuing with the example given above, there would be two groups: one with 25 countries and one with 23 countries). Indeed, this presents a limitation of this study and implies a potential area for future inquiry, which will be discussed in the conclusions.

4. Discussion

Based on the interpretation method described in the previous section, the results can be summarised as follows.

Among both high-GDP and middle-GDP countries, greater intergenerational mobility in education is associated with higher levels of education in general, lower income inequality, higher public spending on education, shorter periods of compulsory education, and a greater emphasis on children's independence.

The only difference between high-GDP and middle-GDP countries appears to be the degree of urbanisation (irrelevant among the former and negatively associated with intergenerational mobility among the latter). This can be explained by considering that cluster D almost exclusively encompasses Latin American and Caribbean countries (Appendix C) that are highly urbanised (on average, 70% of individuals live in cities, a figure second only to North America (78%)). This degree of urbanisation, however, does not protect against significant socio-economic segregation ('favelas' in Brazil and Colombia, 'villa miseria' in Argentina, 'slums' in Mexico, etc.).

Concerning the importance of public spending on education in high-GDP and middle-GDP countries, further evidence can also be extracted from the results. If one orders the clusters according to the level of mobility (thus, B, A, C, D), the same order also applies to the level of public spending on education.

Among low-GDP countries, greater intergenerational mobility in education is associated with a higher degree of urbanisation, higher income inequality, lower public spending on education, longer periods of compulsory education, and a lower emphasis on children's independence.

It is not surprising that even relatively poor countries have a high degree of intergenerational mobility in education (cluster E). As mentioned in the introduction, the creators of the database themselves point out that the relationship between intergenerational mobility and income turns out to be non-monotonic (Van der Weide et al. 2021). This may be due to the fact that most parents are not educated in poor nations (in fact, the level of education

is low in both clusters E and F) and the family of origin does not matter much when the parents are all equally poor. When the level of education and income begins to increase, the gap in opportunities between the children of poor and wealthy families may begin to widen; however, in wealthy economies, it may narrow when the state invests in public education. All these mechanisms seem to be confirmed by the results of the analysis in this article.

On the other hand, it is interesting to note that greater intergenerational mobility seems to be favoured in poor countries by a higher degree of urbanisation and longer periods of compulsory education (which, on the contrary, does not seem important in rich countries). These results are in line with the considerations stated in the introduction (which illustrates the possible mechanisms in place). Equally, the importance given to children's independence is more important to improving intergenerational mobility in advanced economies, in which households have additional resources they may devote to private investment in education. The same line of reasoning can be applied to the fact that greater income inequality is associated with lower intergenerational mobility in rich countries (in which household income can have a greater influence on children's outcomes through private investment), which is not the case in poor countries.

5. Conclusions

In drawing conclusions, it is important to emphasise that the analysis performed in this study allows us to derive an association between the mean values of the variables in the different clusters but does not allow us to make any statements regarding causality. In the case of income inequality, the dilemma is obvious: Does lower intergenerational mobility favour the persistence of income inequality? Or does high income inequality undermine the equality of opportunity? The topic is widely debated, but on balance it does not seem far-fetched to assume that the risk of a vicious circle is present in medium- and high-GDP countries. This vicious circle can, however, be broken by public investment in education as the results of this analysis show. In this case, the causality, although not formally demonstrated, seems clear: it is easier to imagine that a higher level of public spending encourages greater intergenerational mobility rather than the other way around.

The direction of the relationship between intergenerational mobility and the importance given to children's independence is also ambiguous. Does fostering children's independence promote mobility (because rich families do not indulge their children too much)? Or do I want my child to be independent because I know he or she has the opportunity?

While bearing in mind the fact that the lack of identification of causal links is a strong limitation, based on the (almost descriptive) evidence obtained it is possible to extract some indications in terms of policies that could help to improve intergenerational mobility. When differentiating between less- and more-developed countries (a distinction that is one of the key elements of this study), the following conclusions can be drawn:

- For less-developed countries, public policies should limit the effects of physical, social, and economic segregation and increase the number of years of compulsory education. Public spending on education (as a percentage of GDP) does not seem to be directly relevant (in line with some considerations stated in the literature and pointed out in the introduction). However, it becomes relevant again through the above two channels (less segregation and more compulsory education) because both would still require public intervention at their own cost;
- For more-developed countries, increasing the general level of education and increasing the general level of public spending on education are key elements and, of course, can go hand in hand. Income redistribution policies can also be important (they act through the channel of redistributing income-related opportunities).

The importance given to children's independence is a cultural trait that seems to play a role in developed countries; however, it does not seem to be an area where direct public intervention is possible or appropriate. Indirectly, policies that increase equality of opportunity could be helpful so that, over time, parents feel able to rely on their children's

independence. It is clear, however, that this discourse opens up considerations of virtuous or vicious circles as pointed out above.

The greatest utility of this analysis lies in the identification of 'patterns' by grouping countries according to the level of intergenerational education mobility (or 'immobility') and other variables that may be associated with it. This can help us understand the phenomenon of intergenerational mobility in education. The results are in line with those highlighted in the literature and thus support and strengthen them, including by highlighting the non-monotonicity of mechanisms and relationships among countries with different levels of development. This analysis also emphasises the 'equalising' role played by public spending on education in middle- and high-income countries. Finally, an element new to the empirical literature was introduced into this analysis: the explicit consideration of the role played by cultural attitudes with respect to children's independence.

In terms of directions for future research, two aspects can be highlighted. The first is the need to deepen the causal relationships between intergenerational mobility and the factors that may influence it while differentiating between more- and less-developed countries. The second is the need to take cultural aspects into account, both theoretically and empirically.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data used in the article were all taken from free, publicly accessible databases. In addition, the data will be made available by the author on request.

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

Appendix A

Table A1. Countries included in the analysis, sorted according to the value of MOB_EDU ('Expected rank of a child whose parents rank in the bottom half of the education distribution').

Country	MOB_EDU	Country	MOB_EDU	Country	MOB_EDU
Cyprus	49.8	Portugal	40.4	Iran	37.2
Denmark	47.1	Hungary	40.3	Malaysia	37.1
Maldives	46.7	Serbia	40.2	Austria	37.1
Philippines	46.1	Switzerland	40.0	Azerbaijan	37.0
Israel	45.6	Ukraine	39.9	Uruguay	37.0
Russian Fed.	45.5	Kyrgyz Rep.	39.8	Greece	36.9
Zambia	45.1	Egypt	39.8	Iraq	36.8
Iceland	44.9	Korea, Rep.	39.6	Guatemala	36.7
Slovenia	44.4	Sweden	39.6	El Salvador	36.6
Germany	43.8	South Africa	39.1	Kenya	36.6
Slovak Rep.	43.1	Thailand	39.0	China	36.5
Finland	42.8	Belarus	38.9	Brazil	36.3
Belgium	42.8	Italy	38.8	Mexico	36.0
Dominican Rep.	42.5	Poland	38.4	Bangladesh	35.8
Turkey	42.3	Vietnam	38.4	Chile	35.7
Uzbekistan	41.9	Bulgaria	38.4	Ireland	35.6

Table A1. *Cont.*

Country	MOB_EDU	Country	MOB_EDU	Country	MOB_EDU
United Kingdom	41.9	Armenia	38.3	Peru	35.6
Albania	41.9	Moldova	38.3	United States	35.5
France	41.8	Australia	38.2	Romania	35.3
Estonia	41.7	Croatia	38.2	Pakistan	35.0
Japan	41.7	Argentina	38.0	Ecuador	34.9
The Netherlands	41.3	Spain	37.8	Nigeria	34.4
Latvia	41.3	Georgia	37.5	Colombia	34.3
Norway	41.0	Czech Rep.	37.4	Bolivia	34.2
Indonesia	41.0	Kazakhstan	37.4	Lebanon	33.5
Jordan	40.9	Mongolia	37.3	Ghana	32.7
Canada	40.8	Tunisia	37.3		
Lithuania	40.8	India	37.3		

Appendix B

Additional technical information on the variables relevant to the reproducibility of the analysis.

Codes of the indicators in the World Bank World Development Indicator (WDI) database (https://datacatalog.worldbank.org/search/dataset/0037712, accessed on 20 July 2023):

- GDP per capita (Constant 2017 Int. Dollars): NY.GDP.PCAPP.PP.KD
- Educational attainment (% of the population 25+ who at least completed upper secondary school): SE.SEC.CUAT.UP.ZS
- Urban population (% of tot. population): SP.URB.TOTL.IN.ZS
- Gini Index: SI.POV.GINI
- Government expenditure on education (% GDP): SE.XPD.TOTL.GD.ZS
- Compulsory education (years): SE.COM.DURS

The variable 'Share of parents indicating 'Children's Independence' as an important quality in children (%)' comes from the *European Values Study—World Values Study Integrated survey 1981–2021* (https://www.worldvaluessurvey.org/WVSEVStrend.jsp, accessed on 27 July 2023).

The Integrated Values Survey (IVS) dataset 1981–2021 can be constructed by merging the EVS Trend File 1981–2017 (doi:10.4232/1.13736) and the WVS time series 1981–2021 dataset (doi:10.14281/18241.15). It is based on the Common EVS/WVS Dictionary (2021).

It is also possible to find the IVS merge syntax for Stata at following link: https://www.worldvaluessurvey.org/WVSEVStrend.jsp, accessed on 20 July 2023.

The code of the variable used to derive the percentage of parents who value their children's independence is A029 'Important child qualities: independence' (see the IVS Common EVS–WVS dictionary at https://www.worldvaluessurvey.org/WVSEVStrend.jsp for further details). The individual weight used in order to obtain the country average (for each available year) is the variable whose code is S017.

The variable regarding intergenerational mobility in education ('Expected rank of a child whose parents rank in the bottom half of the education distribution') comes from the World Bank Global Database on Intergenerational Mobility (GDIM): https://datacatalog.worldbank.org/search/dataset/0050771/Global-Database-on-Intergenerational-Mobility, accessed on 17 July 2023.

In that database, the variable name/code is 'MU050_randomtiebreak'.

Appendix C

Countries falling within the clusters identified in Table 2.

Cluster A (High GDP, lower intergenerational mobility): Australia, Belgium, Canada, France, Germany, Hungary, Ireland, Italy, Latvia, The Netherlands, Spain, Switzerland, United Kingdom, United States.

Cluster B (High GDP, higher intergenerational mobility): Austria, Cyprus, Denmark, Finland, Iceland, Israel, Japan, Lithuania, Norway, Slovenia, Sweden.

Cluster C (Middle GDP, higher intergenerational mobility): Armenia, Azerbaijan, Belarus, Bulgaria, Croatia, Czech Republic, Estonia, Georgia, Greece, Jordan, Kazakhstan, Korea, Rep., Kyrgyz Republic, Moldova, Mongolia, Poland, Romania, Russian Federation, Serbia, Slovak Republic, Ukraine, Uzbekistan.

Cluster D (Middle GDP, lower intergenerational mobility): Argentina, Bolivia, Brazil, Chile, Colombia, Ecuador, El Salvador, Guatemala, Mexico, Peru, South Africa, Uruguay.

Cluster E (Low GDP, higher intergenerational mobility): Dominican Republic, Maldives, Philippines, Turkey, Zambia.

Cluster F (Low GDP, lower intergenerational mobility): Albania, Bangladesh, China, Egypt, Arab Rep., Ghana, India, Indonesia, Iran, Iraq, Kenya, Lebanon, Malaysia, Nigeria, Pakistan, Portugal, Thailand, Tunisia, Vietnam.

References

Balcazar, Carlos Felipe, Ambar Narayan, and Sailesh Tiwari. 2015. *Born with a Silver Spoon: Inequality in Educational Achievement across the World*. Policy Research Working Paper 7152. Washington, DC: World Bank.

Becker, Gary S., and Nigel Tomes. 1979. An equilibrium theory of the distribution of income and intergenerational mobility. *Journal of Political Economy* 87: 1153–89. Available online: https://www.jstor.org/stable/1833328 (accessed on 3 February 2024). [CrossRef]

Becker, Gary S., Scott Duke Kominers, Kevin M. Murphy, and Jörg L. Spenkuch. 2018. A Theory of Intergenerational Mobility. *Journal of Political Economy* 126: S7–S25. [CrossRef]

Campbell, Mary, R. Haveman, G. Sandefur, and Barbara Wolfe. 2005. Economic inequality and educational attainment across a generation. *Focus* 23: 11–15.

Chetty, Raj, Nathaniel Hendren, Patrick Kline, and Emmanuel Saez. 2014. Where is the land of opportunity? The geography of intergenerational mobility in the United States. *The Quarterly Journal of Economics* 129: 1553–623. [CrossRef]

Connolly, Marie, Miles Corak, and Catherine Haeck. 2019. Intergenerational Mobility between and within Canada and the United States. *The Journal of Labor Economics* 37: S595–S641. [CrossRef]

Corak, Miles. 2013. Income Inequality, Equality of Opportunity, and Intergenerational Mobility. *Journal of Economic Perspectives* 27: 79–102. [CrossRef]

Corak, Miles. 2021. The Canadian Geography of Intergenerational Income Mobility. *The Economic Journal* 130: 2134–74. [CrossRef]

D'Addio, Anna Cristina. 2007. *Intergenerational Transmission of Disadvantage: Mobility or Immobility across Generations?* OECD Social, Employment and Migration Working Papers, No. 52. Paris: OECD Publishing. [CrossRef]

Davies, James Byron, Jie Zhang, and Jinli Zeng. 2005. Intergenerational Mobility under Private vs. Public Education. *Scandinavian Journal of Economics* 107: 399–417. [CrossRef]

Duncan, Greg J., and Richard J. Murnane. 2012. *Whither Opportunity? Rising Inequality, Schools, and Children's Life Changes*. New York, NY: Russell Sage Foundation.

Durlauf, Steven N., and Ananth Seshadri. 2018. Understading the Great Gatsby Curve. In *NBER Macroeconomics Annual 2017*. Edited by Martin S. Eichenbaum and Jonathan Parker. Chicago: University of Chicago Press, vol. 32.

Feinstein, Leon, Ricardo Sebates, Tashweka Anderson, Annik Sorhaindo, and Cathie Hammond. 2006. *The Effects of Education on Health: Concepts, Evidence and Policy Implications*. A Review for the OECD Centre for Educational Research and Innovation (CERI). Paris: CERI.

GDIM (Global Database on Intergenerational Mobility). 2023. *Development Research Group, World Bank*. Washington, DC: World Bank Group. Available online: https://datacatalog.worldbank.org/search/dataset/0050771/Global-Database-on-Intergenerational-Mobility (accessed on 17 July 2023).

Glomm, Gerhard, and B. Ravikumar. 1992. Public versus Private Investment in Human Capital: Endogenous Growth and Income Inequality. *Journal of Political Economy* 100: 818–34. [CrossRef]

Hastie, Trevor, Robert Tibshirani, and Jerome Friedman. 2009. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*, 2nd ed. New York: Springer.

Herrington, Cristopher Michael. 2015. Public education financing, earnings inequality, and intergenerational mobility. *Review of Economic Dynamics* 18: 822–42. [CrossRef]

Jerrim, John, and Lindsey Macmillan. 2015. Income Inequality, Intergenerational Mobility, and the Great Gatsby Curve: Is Education the Key? *Social Forces* 94: 505–33. [CrossRef]

Kaufmann, Leonard, and Peter J. Rousseeuw. 1990. *Finding Groups in Data: An Introduction to Cluster Analysis*. New York, NY: Wiley.
Kearney, Melissa S., and Phillip B. Levine. 2016. Income Inequality, Social Mobility, and the Decision to Drop Out of High School. *Brookings Papers on Economic Activity Spring* 333–96. Available online: http://www.nber.org/papers/w20195 (accessed on 20 February 2024).
Lee, Sang Yoon (Tim), and Ananth Seshadri. 2019. On the Intergenerational Transmission of Economic Status. *Journal of Political Economy* 127: 855–921. [CrossRef] [PubMed]
Leone, Tharcisio. 2019. *Intergenerational Mobility in Education: Estimates of the Worldwide Variation*. UNRISD Occasional Paper—Overcoming Inequalities in a Fractured World: Between Elite Power and Social Mobilization, No. 2. Geneva: United Nations Research Institute for Social Development (UNRISD). ISBN 978-92-9085-103-5.
Lloyd-Ellis, Huw. 2000. Public Education, Occupational Choice, and the Growth-Inequality Relationship. *International Economic Review* 41: 171–202. Available online: http://www.jstor.org/stable/2648828 (accessed on 25 February 2024). [CrossRef]
Makles, Anna. 2012. Stata tip 110: How to get the optimal k-means cluster solution. *The Stata Journal* 12: 347–51. [CrossRef]
Mayer, Susan E. 1997. *What Money Can't Buy: Family Income and Children's Life Chances*. Cambridge: Harvard University Press.
Mayer, Susan E., and Leonard M. Lopoo. 2008. Government Spending and Intergenerational Mobility. *Journal of Public Economics* 92: 139–58. [CrossRef]
McLanahan, Sara, ed. 2020. *How Cultural Factors Shape Economic Outcomes*. Princeton: Princeton and Brookings, The Future of Children, p. 190. Available online: https://www.jstor.org/stable/e27074970 (accessed on 20 March 2024).
Narayan, Ambar, Roy Van der Weide, Alexandru Cojocaru, Christoph Lakner, Silvia Redaelli, Daniel Gerszon Mahler, Rakesh Gupta N. Ramasubbaiah, and Stefan Thewissen. 2018. *Fair Progress?: Economic Mobility Across Generations Around the World*. Washington, DC: World Bank Publications.
Neidhöfer, Guido, Joaquin Serrano, and Leonardo Gasparini. 2018. Educational Inequality and Intergenerational Mobility in Latin America: A New Database. *Journal of Development Economics* 134: 329–49. [CrossRef]
Piketty, Thomas. 2000. Chapter 8: Theories of persistent inequality and intergenerational mobility. In *Handbook of Income Distribution*. Amsterdam: Elsevier, vol. 1, pp. 429–76.
Reardon, Sean. 2011. The widening academic achievement gap between rich and poor: New evidence and possible explanations. In *Whither Opportunity? Rising Inequality, Schools, and Children's Life Changes*. Edited by Greg J. Duncan and Richard J. Murnane. New York: Russell Sage Foundation, pp. 91–115.
Reddy, Amarender, and Vikas Singh. 2021. Do parents' education predict child education and occupation? *IASSI-Quarterly* 40: 644–63.
Restuccia, Diego, and Carlos Urrutia. 2004. Intergenerational Persistence of Earnings: The Role of Early and College Education. *American Economic Review* 94: 1354–78. [CrossRef]
Sen, Amartya. 1980. Equality of What? In *Tanner Lectures on Human Values*. Edited by Sterling M. McMurrin. Cambridge: Cambridge University Press, vol. 1.
Solon, Gary. 2004. A Model of Intergenerational Mobility Variation over Time and Place. In *Generational Income Mobility in North America and Europe*. Edited by M. Corak. Cambridge: Cambridge University Press, chp. 2, pp. 38–47.
Staffolani, Stefano, and Enzo Valentini. 2007. Bequest Taxation and Efficient Allocation of Talents. *Economic Modelling* 24: 648–72. [CrossRef]
Stuhler, Jan. 2018. *A Review of Intergenerational Mobility and Its Drivers*. JRC Research Reports, JRC112247. Seville: European Commission, Joint Research Centre (Seville site).
Van der Weide, Roy, Christoph Lakner, Daniel Gerszon Mahler, Ambar Narayan, and Rakesh Ramasubbaiah. 2021. *Intergenerational Mobility around the World*. Policy Research Working Paper; No. 9707. Washington, DC: World Bank.
Van der Weide, Roy, Christoph Lakner, Daniel Gerszon Mahler, Ambar Narayan, and Rakesh Gupta. 2023. Intergenerational mobility around the world: A new database. *Journal of Development Economics* 166: 103167. [CrossRef]

economies

MDPI

Article

The Impact of the Great Recession on Well-Being across Europe Ten Years On: A Cluster Analysis

Elisabetta Croci Angelini [1,*], Francesco Farina [2] and Silvia Sorana [3]

[1] Center for Interdisciplinary Studies in Economics, Psychology and Social Sciences (CISEPS), University of Milano-Bicocca, 20126 Milan, Italy
[2] Centre for Investigation and Modelling of Experimental Observations (CIMEO), Sapienza University of Rome, 00161 Rome, Italy
[3] COOSS Marche Onlus, 60121 Ancona, Italy
* Correspondence: elisabettacroci@gmail.com

Abstract: To evaluate variations in the well-being dimensions of European citizens, we rely upon Principal Component Analysis methodology, whereby a large set of interrelated indicators are reduced to a small number of aggregate synthetic variables. We find that the 2008 crisis impinged differently on the various dimensions of well-being. The evolution of the indicators has affected different clusters of countries in various ways. Most importantly, we observe that there has been a shift of the principal component from the poor in terms of material deprivation to the risk of poverty for the worsening conditions in the labor market.

Keywords: European Union; Great Recession; well-being; Principal Component Analysis

Citation: Croci Angelini, Elisabetta, Francesco Farina, and Silvia Sorana. 2024. The Impact of the Great Recession on Well-Being across Europe Ten Years On: A Cluster Analysis. *Economies* 12: 115. https://doi.org/10.3390/economies12050115

Academic Editor: Bruce Morley

Received: 1 April 2024
Revised: 30 April 2024
Accepted: 7 May 2024
Published: 10 May 2024

1. Introduction

The Great Recession had widespread consequences. During the 2008 financial crisis the interplay between the mutual exposure of banks and governments to the other party's insolvency risk greatly distressed the balance sheets of the banks, finally leading to a credit crunch. A severe recession in advanced countries, with rising unemployment and negative growth rates, caused a lower demand level. In Europe, diminishing earnings for households and declining profits for firms coupled with the slackened functioning of automatic stabilizers posited the Eurozone GDP dynamics on a lower path (De Grauwe and Ji 2013). To counter rising public deficit and public debts with respect to GDP, austerity policies were implemented through restrictive impulses of fiscal policy.

The less efficient countries participating in the European Monetary Union have been exposed to the divergent impact of the common monetary policy and of the common fiscal rules. Due to the austerity policies meant to recover competitiveness, real devaluation ensued, with both lower employment rates and substantial wage cuts. Rocketing risk premia increased the spread of the Eurozone's sovereign bonds *vis-à-vis* the German 10-year bund, with particularly high hikes in peripheral countries due to a contagion effect triggered by the partial default of the Greek public debt (Croci Angelini et al. 2016).

The burden of the labor market adjustment has disproportionally fallen on the low-skilled labor force through lower job protection and lower pay, and to a larger extent on non-standard jobs. The widening top–bottom income inequality has been due more to increasing unemployment than to an enlarging distance between the bottom section of the wage distribution and the average wage, although in a few countries the reduction in earnings—larger for high-income than for low-income households—has slightly reduced income inequality (ILO 2015). Cuts in social expenses reduced the public provision of both monetary transfers and in-kind services. Not only were pensions and unemployment benefits reduced, but the degree of coverage, targeting, and generosity shrank too. These developments impinged not only on the earnings dimension, but also on quality of life, as

shown by relevant variations in the main non-monetary well-being dimensions in different countries. Enduring consequences for living conditions were registered in crucial dimensions of well-being, such as health, education, and also social inclusion (Jenkins et al. 2013). In 1985, the Council of Ministers of the European Union defined the "poor" as: "the persons whose resources (material, cultural and social) are so limited as to exclude them from the minimum acceptable way of life in the Member State to which they belong" (EU Council of Ministers 1985). As an effect of the lowering disposable income caused by countercyclical fiscal impulses and more flexible labor markets, the vulnerability of low-skilled workers and the precarious conditions of the labor force with part-time jobs expanded.

Empirical evidence shows that in two-thirds of OECD countries, income inequality has been growing hand in hand with relative poverty; in all of them the risk of poverty, and in some of them also the intensity of poverty, has been soaring (Atkinson et al. 2010). The percentage of materially deprived people ranged from 3% in Luxembourg and 6% in Sweden and the Netherlands up to 45% in Latvia. These distances were much wider than the dispersion of poverty risk, ranging from 10% to 21% (Fusco et al. 2010, pp. 137–38). However, the heterogeneity in living conditions across European countries extends beyond the income-based indicators of poverty and inequality. An Index of Economic Well-being, constructed by incorporating information on consumption, wealth, inequality, and economic insecurity for the OECD countries, shows that the economic crisis more negatively affected well-being in the Peripheral Eurozone, *vis-à-vis* the other countries in the sample.

Low-income people and the poor were also disproportionately affected by the intensification in non-material deprivation. The risk of poverty soared mainly in the sub-group of involuntarily part-time workers (Horemans et al. 2015). While the most dramatic fall in income was seen in the poor in Greece, empirical evidence shows that in many countries low-income groups and the poor were more severely hit during the Great Recession in terms of socio-economic attributes (Whelan and Maître 2012). Between 2008 and 2012 the fall in equivalized household income was very large in Iceland (40%), Greece (30%), and Ireland (20%), and to a lesser extent in the United Kingdom, Spain, and Portugal, where reductions ranged from 13% to 10%. To convey the well-being distance across people concerning material resources, the at-risk-of-poverty indicator has been employed. In 2008, 19 million people were living in severely materially deprived households in Europe; 17 million individuals aged 0–59 were living in jobless households; 49.6 million were living in households at risk of poverty, but were neither jobless nor severely materially deprived; 40 million were living in jobless households and/or materially deprived even if not at risk of poverty; whereas 6.9 million were living in jobless households, at risk of poverty, and severely materially deprived. "The rate for the 12 'new' Member States (NMS12) was 17.3 per cent, a little but not much higher than for EU-15 with a rate of 16.4 per cent. It is certainly not the case that those at risk of poverty on the EU definition are mostly to be found in the New Member States: of the 80+ million at risk of poverty in EU-27, 64 million are to be found in the EU15. In Germany, alone, there are $12\frac{1}{2}$ million; in the United Kingdom $11\frac{1}{2}$ million; in Italy 11 million; and France and Spain together account for a further 17 million. In the largest New Member State, Poland, the number of people at risk of poverty is about $11\frac{1}{2}$ million" (Atkinson and Marlier 2010, p. 106).

The Europe 2020 Agenda, adopted by the European Union (European Commission 2010), pointed at making substantive progress, among other goals, in the promotion of social inclusion. The objectives set in Lisbon in 2000 were neither entirely accomplished by all member states, nor by the EU as a whole (Grimaccia 2021). Also due to the COVID-19 pandemic, the "strategy for a smart, inclusive and sustainable growth" did not deliver its promises. The 2020 target in the area of poverty and social exclusion was defined on the basis of three indicators: (1) the number of people considered 'at-risk-of-poverty' according to the EU definition, where the poverty risk threshold set at 60% of the national household equivalized median income; (2) the number of materially deprived people; and (3) the number of people aged 0–59 living in 'jobless' households, where no member aged 18–59 is working, or where members aged 18–59 have, on average, very limited work attachment.

The target was set to reduce by 25% the number of Europeans living below national poverty lines, by lifting around 20 million people out of poverty. The problem of fulfilling the objectives for social inclusion set in the Europe 2020 Agenda was considerably exacerbated (Atkinson and Marlier 2010).

The Great Recession hit European countries differently by increasing the risk of poverty and inequality and, within those countries, hit individual households in terms of material and non-material deprivations, which include health and education as well as other dimensions relevant for the quality of life—a concept whose content is still debated, but on whose multidimensionality there is no doubt.

Our paper investigates how the real devaluation ensuing the Great Recession affected the multidimensional well-being in European countries through Principal Component Analysis (PCA). The main idea is to summarize in a few major points the differences, if any, encountered by households after the crisis, how their well-being was affected at that time, and whether or not they found it difficult to recover. While there is no lack of studies addressing single issues at the country level, especially from a macroeconomic point of view, not many are based on microdata and explore household behavior in a multidimensional framework by applying PCA.

The paper is organized as follows: Section 2 reviews some empirical literature meant to frame the setting and where the relevance of multidimensional well-being has been addressed; Section 3 discusses the methodological choice of analyzing the impact of the Great Recession on European well-being through PCA, also compared with other methods. The results are presented in Section 4 and discussed in Section 5 where a comparison of our results with previous research is provided. The 2008 crisis impinged differently on the various dimensions of well-being. All in all, our findings very much differentiate depending on the indicators and on the different groups of countries. Section 6 concludes the paper.

2. Literature Review

The concern about income distribution and social exclusion in private households in the EU compellingly emerged at the turn of the century. The European Union Statistics on Income and Living Conditions (EU-SILC) dataset was established in the framework of the Open Method of Coordination within the Programme of Community Action meant to encourage cooperation between Member States to counteract social exclusion. It covers European countries, not necessarily members of the EU, and aims at issuing comparable statistics through an integrated design in this area of inquiry. A very wide literature appeared focusing on concepts, measurement, and evaluations of different aspects of inequality. By their very nature, socio-economic phenomena are the joint product of a variety of micro-economic characteristics (e.g., individual material deprivation) and/or macroeconomic conditions (e.g., an underemployment equilibrium with jobless households and/or individuals at high risk of poverty) impinging on the well-being of society at large while interrelations across the most relevant variables are difficult to disentangle.

Following the "capability approach" (Sen 1985), in recent decades the research on well-being has turned towards multidimensionality. Indeed, quality of life is a multifaceted concept (Nussbaum and Sen 1993) to be achieved through a series of functionings, consisting of opportunities in terms of personal capacities. The empirical strand of literature on social indicators has increased enormously in the last three decades, suggesting that, to obtain a comprehensive evaluation of well-being, more dimensions need be added to the standard monetary dimension (Nolan and Whelan 2014).

Divergent per capita GDP across the EU member states, and the dispersion of socio-economic status within the population, are bound to impinge on health conditions (Crimmins et al. 2009). According to EU-SILC data (which refer to self-perceived health status, longstanding illness or disability, unmet medical and dental treatments, and limitations on daily activity), health limitations impaired activity levels between up to

around one-fifth in Cyprus, Poland, Sweden, and the United Kingdom, and over one-third in Estonia, Finland, and Latvia (Hernández-Quevedo et al. 2010).

An analysis conducted by the EUROMOD team shows that the key factor in protecting a household from a drop in income is the presence of more than one single member of the household earning an income (Figari et al. 2010). Two individuals with the same income can have very different living standards if the resources available to each of them differ because of different national provision of public transfers (Fusco et al. 2010). Welfare institutions were crucial in the reduction of the risk of poverty within the European Union, ranging from under 15% to over 60%, with an average value of 38% (Whelan et al. 2014).

While the evolution of jobs, especially for low-income households, is certainly relevant in the cross-country comparison of well-being in Europe before and after the crisis, the effect of the switch towards a more flexible labor market is difficult to assess. During the Great Recession, unemployment in the OECD labor force rose from 6.6% to 8%, with youth unemployment doubling on average, and reaching a peak of 50% in Greece and Spain (OECD 2015). Recent estimates convey the message that a possible positive impact on the employment rate very much depends on the initial degree of rigidity and the mix of institutional reforms (Sologon and O'Donoghue 2014).

As for accommodation, an OECD Statistical Brief reports that housing prices, along with the savings ratio, represent a key driver of the level of household wealth, as the positive correlation between the median net wealth of households and the annual real growth rate of house prices is strong in the long run (Murtin and Mira d'Ercole 2015, p. 4). The relationship between income and wealth is also influenced in Europe by the varying impact across clusters of EU countries of the different forms of housing tenure (Kemeny 2001; Croci Angelini 2015). To compare the standard of living of owner-occupiers and tenants, the method adopted by Eurostat consists in the "imputation" of a rent to owners (having subtracted the actual housing costs). Overall, the adjustment performed by means of the inclusion of imputed rent reduces the degree of income inequality. In particular, the at-risk-of-poverty rate would fall by five percentage points in Ireland and the United Kingdom, four in Estonia and Spain, and more than two in Belgium, Greece, Latvia, and Portugal (Sauli and Törmälehto 2010). Although home-ownership disproportionately affects the well-being of high-income versus low-income households, the impact of house property on a household's financial balances is heterogeneous. On the one hand, the owner-occupier benefits from a higher income, as he gains a hidden rent (corresponding to the saved rent, which would have been paid to a landlord); also, a retired worker living in his own flat, but on the brink of poverty because of a low pension, could use his house to obtain a loan from a bank so as to improve a poor lifestyle. On the other hand, the loss in household equivalized income during a recession is countered in some countries by offering home-ownership as the collateral to borrowing, while in other countries the mortgage associated with home-ownership may worsen already distressed household finances, mainly depending on the income level of households and the national percentage of home-ownership (Sierminska 2012). In some EU countries the indicators for housing conditions were found to be highly correlated with income, while the indicators of material deprivation usually present a stronger relationship with income than with housing conditions, mainly as an effect of financial stress (Nolan and Whelan 2010). Furthermore, the sudden fall in short-term income impacts everyday life, and a declining long-term income counts more for housing conditions and the social environment; similarly, the degree of deprivation is higher for financial distress than for worsening social environment and housing conditions (Fusco et al. 2010).

In the research effort aimed at evaluating multidimensional well-being (MWB), a methodological issue has to be tackled. The dimension-by-dimension approach—a "large and eclectic dashboard" (Stiglitz et al. 2009), or a "portfolio of indicators" (Atkinson et al. 2002)—aims at preserving the information on the interpersonal dispersion of well-being in each dimension. A synthetic indicator may summarize the overall well-being at the cost of ignoring possible interactions across dimensions. On this issue, the empirical evidence

stemming from the EU-SILC database is unclear. On the one hand, the estimate of three indicators—being at risk of poverty, living in a jobless household, and suffering from material deprivation—shows that one-third of the individuals are "disadvantaged" in more than one dimension (Atkinson et al. 2010, p. 127). On the other hand, everywhere there is a low correlation between the income level and the level of deprivation; in particular, in some countries a high level of deprivation is associated with a low level of poverty (Atkinson and Marlier 2010). To compute an index for each dimension avoids two critical issues: the normative evaluation of the weight to be attributed to each dimension, and the assessment of the degree of substitutability among them (Decancq and Lugo 2013).

3. Methodology

Multidimensional well-being (MWB) indices seek the impact on well-being stemming from the mutual reinforcement of conditions often characterized by a high degree of complementarity. The variables they rely upon seldom enjoy orthogonality, a characteristic the lack of which hinders many quantitative analyses. The inputs face the problems of identifying the relevant dimensions, find indicators able to describe them, and aggregate the indicators into a single figure meant to aptly describe the multidimensional phenomenon. From a theoretical point of view, several characteristics are needed, a requirement that has been coped with by axiomatic methodologies (Weymark 2006). Empirically, they are often based on surveys, such as the European Quality of Life Survey (EQLS), where questions are posed by Eurofound to thousands of selected individuals. The queries are obviously designed to fit the purpose of the survey, yet independent researchers may use the data for their own investigations.

To compare well-being in European countries before and after the Great Recession we use data from the EU-SILC dataset for the years 2007 and 2012. The data—covering 26 countries, among which 24 belong to the European Union and two are non-EU countries (Norway and Iceland)—are complete, i.e., all relevant variables exist for both years and all countries. Although Bulgaria, Croatia, Malta, and Romania are EU members today, they have been excluded for incomplete availability of data. Our units of analysis are these 26 European countries: for each of them the dataset includes several thousand entries, based on both households and individuals, from which each country's information is calculated.

The Principal Component Analysis (PCA) searches for the unknown factors which are at the roots of the well-being outcomes. This methodology consists of the computation of mutually orthogonal principal components (through the linear combinations of the original variables, i.e., the different indicators considered for each dimension). A large number of initial, possibly correlated, indicators are transformed into mutually uncorrelated linear combinations. The principal components are extracted with the aim of identifying hidden, unobservable variables able to explain a major portion of the variance. Hence, dispersed information about each individual entry is concentrated in principal components, each one summarizing the information conveyed by a larger set of indicators. The construction of composite indices from individual indicators helps in comparing across time and space the performance of a unit based on a large amount of information (Freudenberg 2003). The PCA is a reliable method meant to overcome the trade-off between comprehensiveness (which compresses the variety of dimensions of life into a synthetic index) and meaning (whereby the focus on the impact of the crisis on well-being prompts preserving the distinct short-term evolutionary path of well-being in each dimension) and so helps in weighting performances and devising policies (Nardo et al. 2008). To evaluate unobservable variables such as well-being or quality of life, an alternative method is the fuzzy set approach (Betti 2016) where the methodological focus is on the appropriate weights, while another method is by axiomatic measurement, which keeps a desirable decomposition characteristic and was proposed in a previous paper (Croci Angelini and Michelangeli 2012).

Our methodological choice in favor of PCA was based on the reduction of variables aimed at understanding the structure underlying a large list of interrelated indicators in order to reduce them to a small number of aggregate synthetic variables. The search

for latent variables is particularly suitable for dealing with the large amounts of data characterizing a socio-economic survey on income, quality of life, and living conditions such as those provided by EU-SILC, the dataset we rely upon.

By sorting 68 variables from the EQLS dataset, Betti (2016) identifies eight relevant groups, ranging from quality of relations to subjective well-being, and including health and housing quality. On the contrary, we consider the five dimensions we regard as most relevant for our research. With the three most investigated variables in the analysis of multidimensional well-being (income, education, and health), this paper addresses employment and accommodation. These two additional dimensions are relevant for the lack of territorial homogeneity in Europe, where across-country mobility is more demanding and risky, due to institutional as well as cultural differences.

EU-SILC includes both subjective and objective queries. While we aim to keep as much information as possible from the dataset, our strategy is to rely upon objective information only and to check its coherence with subjective evaluations, when available. This check is important especially when the eligible answers are not dichotomous and allow either a scale of preferences, or provide reasons to support the answer. A case in point is when the answer "Yes" differentiates between intensities, or the answer "No" distinguishes between preference (those who do not want the item) and feasibility (those who cannot afford it).

Following the OECD, both material and non-material sources of well-being have been considered. However, our classification does not exactly reflect the same sub-sets of dimensions separating the more observable material living conditions from the (somewhat more) subjective quality of life. The OECD framework for the assessment of well-being considers outcomes and their distribution across the population achieved in two broad domains: material living conditions (i.e., income and wealth, jobs and earnings, and housing conditions) and quality of life (i.e., health status, work-life balance, education and skills, social connections, civic engagement and governance. environmental quality, personal security, and subjective well-being).

As for the material sources of well-being, monetary and non-monetary dimensions have been chosen, relevant at the household level and quantified through three observable indicators:

1. Monetary income, which singles out individuals with an equivalized disposable income below 60% of their national median and is assessed through the dichotomous variable "at-risk-of-poverty". Equivalized disposable income corresponds to the total household income after social transfers, available for spending or saving, divided by the number of household members, converted into adults, according to the OECD modified equivalence scale.

2. Material deprivation, which refers to difficulties in everyday life stemming from lack of key provisions. The Eurostat definition covers a set of sub-indicators of economic strain, related to a household's inability to cope with a number of items deemed essential for a decent life. According to this definition, materially deprived individuals are those who cannot afford at least three out of the following list of nine items: (1) paying for unexpected financial expenses, (2) keeping their homes adequately warm, (3) eating meat, fish, or a protein equivalent every second day, (4) enjoying a week's holiday away from home, (5) having a car, (6) having a washing machine, (7) having a color TV, or (8) having a telephone, and (9) paying mortgage, rent, utility bills, as well as hire purchase installments or other loan payments without delay.

3. Housing deprivation, which looks at poorly-comfortable dwellings. Following the Eurostat definition, to be severely deprived, a household should live in an overcrowded dwelling, which also suffers from at least one of the following deficiencies: (1) being too dark, (2) having a leaking roof, damp walls/floors/foundation, or rot in the window frames or floor, and (3) having neither a bath, nor a shower, nor an indoor toilet. The Eurostat methodology for the calculation of the overcrowding variable employs auxiliary variables referring to the number, age, and gender of the household members.

As for non-material sources of well-being, three dimensions, relevant at the personal level, have been considered and measured by indicators as objective as possible:

1. Health, which collects all objective information available in EU-SILC on this dimension. The variables included are: (1) suffering from a chronic illness, (2) unmet medical treatment, (3) unmet dental treatment, and (4) activities limited by impaired health conditions. Information about personal general health assessments was deemed subjective and was overlooked.

2. Education, which measures young individuals not in employment, education, or training (NEETs)—a condition which points at the mismatch between jobs and education. This variable was computed by selecting the individuals aged 16–29 whose economic activity was non-existent and who participated neither in education nor in training.

3. Economic activity, which is summarized by three indicators related to difficulties suffered in the labor market. The variables included are: (1) temporary jobs, (2) unemployment, and (3) under-employment, appraised by work intensity, which refers to households where, during the previous 12 months, their components aged 18–59 worked less than 20% of their total potential.

4. Results

Table 1 shows the correlation matrix of these indicators for the year 2007. The highest correlations are between both deprivations (0.859) and between both unmet treatments (0.824); correlations over 0.5 are also observed between chronic illness and limited activities; for severe housing deprivation and NEETs as well as unmet treatments, NEETs are also associated with unemployment and underemployment. A weak negative correlation is observed between chronic illness and poverty risk (−0.061) and, in turn, the NEETs (−0.138), as well as between unmet dental treatment and unemployment (−0.229), while temporary jobs show limited negative correlations with housing deprivation and all items connected to health.

Table 1. Inter-item correlation matrix for the year 2007.

	1.	**2.**	**3.**	**4.**	**5.**	**6.**	**7.**	**8.**	**9.**	**10.**	**11.**
1. Severe housing deprivation	1										
2. Suffer from chronic illness	0.194	1									
3. Unmet medical treatment	**0.663**	0.194	1								
4. Unmet dental treatment	**0.510**	0.087	**0.824**	1							
5. Activity limited by bad health	0.396	**0.591**	0.352	0.175	1						
6. Poverty risk	0.385	−0.061	0.383	0.381	0.220	1					
7. Extreme material deprivation	**0.859**	0.093	**0.714**	**0.605**	0.347	0.415	1				
8. Unemployment	0.366	0.140	0.046	−0.229	0.398	0.161	0.337	1			
9. Underemployment	0.311	0.033	0.396	0.245	0.240	0.267	**0.507**	0.415	1		
10. Temporary jobs	−0.043	−0.200	−0.057	−0.067	−0.107	0.140	0.046	0.451	0.490	1	
11. NEETs	**0.509**	−0.138	0.300	0.113	0.242	0.427	**0.545**	**0.685**	0.481	0.364	1

Source: Own calculations on EU-SILC 2007 and EU-SILC 2012 dataset. Correlations over 0.5 in bold.

Table 2 shows the correlation matrix of the same indicators for the year 2012. Again, the highest correlations are observed between both deprivations (0.864) and between both unmet treatments (0.705). However, more impressive variations stand out. With respect to 2007, the unemployment correlation increases with poverty risk (from 0.161 to 0.69) and between unemployment and underemployment (from 0.415 to 0.75). NEETs are no longer correlated with any other variable and neither are Temporary jobs. The number of (weak) negative correlation values has also increased, mainly affecting those suffering chronic illness and those employed in temporary jobs.

Table 2. Inter-item correlation matrix for the year 2012.

	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.
1. Severe housing deprivation	1										
2. Suffer from chronic illness	0.061	1									
3. Unmet medical treatment	**0.583**	0.190	1								
4. Unmet dental treatment	0.349	0.094	**0.705**	1							
5. Activity limited by bad health	0.217	**0.545**	0.008	−0.233	1						
6. Poverty risk	0.273	−0.227	0.325	0.294	−0.003	1					
7. Extreme material deprivation	**0.864**	−0.072	**0.515**	0.353	0.088	0.448	1				
8. Unemployment	0.325	−0.221	0.241	0.375	−0.018	**0.690**	0.487	1			
9. Underemployment	0.328	−0.162	0.382	**0.616**	−0.170	**0.567**	0.447	**0.750**	1		
10. Temporary jobs	−0.066	−0.125	0.138	0.167	−0.125	0.299	−0.114	0.479	0.327	1	
11. NEETs	0.186	0.085	0.184	0.105	−0.014	0.245	0.370	0.141	0.191	−0.046	1

Source: Own calculations on EU-SILC 2007 and EU-SILC 2012 dataset. Correlations over 0.5 in bold.

A factor analysis—a method of checking the dimensionality of the scale on the basis of the internal consistency of all items—was performed following the Cronbach's alpha test on reliability. The aim of this exploratory exercise is to single out groups of related variables, each one describing some part of the dimensions we are interested in, but unfit to take part in a regression model. The principal components are extracted from the groups of variables described above.

The results of these tests, performed on the dataset to identify reliability, sampling adequacy, and data suitability, are shown below in Table 3 for both years before and after the outbreak of the Great Recession.

Table 3. Tests on reliability, sampling adequacy, and data stability.

Test on Reliability		**2007**	**2012**
Cronbach's alpha		0.757	0.649
Cronbach's alpha based on standardized items		0.820	0.763
Item number		11	11
Test on Sampling Adequacy		**2007**	**2012**
Measure of sampling adequacy KMO (Keiser Meyer Olkin)		0.709	0.634
Test on Data Suitability		**2007**	**2012**
Bartlett's Test of sphericity	Approximate Chi-square	153.240	139.678
	Degrees of freedom	55	55
	Significance	0.000	0.000

Source: Own calculations on EU-SILC 2007 and EU-SILC 2012 dataset.

The PCA is a reduction technique able to summarize a large set of variables into a smaller number of components so as to seek a pattern resulting from the correlation existing among them. All indicators enter the PCA after being transformed into dichotomous variables. Therefore, information about the causes of difficulties and/or the degree of deprivation when existing at this stage is lost.

The extraction of the components has been performed by employing two methods: the Kaiser eigenvalues (reflecting the components' variance) and the Scree plot test (plotting the eigenvalues and evaluating the shape of the line) so as to seek a good balance between efficiency (i.e., keep the number of components as low as possible so as to describe the results as simply as possible) and completeness (i.e., not to drop any relevant information).

The varimax method, by the rotation matrix of the components, has been applied in order to minimize the number of indicators and keep as much of their information content as possible in the analysis in a situation where many variables were supposed to be cross-correlated (or this possibility could not be ruled out).

Table 4 shows the explanatory power of the three components for the year 2007—and four components for the year 2012—ranked according their decreasing importance. All

Kaiser eigenvalues are higher than 1.0, so that in principle all components are worth being considered in our analysis as the variance they can explain is higher than that explained by the single variables.

Table 4. Total variance explained—rotated factor loadings.

	Total		% Variance		% Cumulative	
Component	**2007**	**2012**	**2007**	**2012**	**2007**	**2012**
1	3.487	2.533	31.700	23.027	31.700	23.027
2	2.674	2.255	24.306	20.498	56.006	43.525
3	1.857	2.048	16.882	18.622	72.888	62.147
4	-	1.597	-	15.520	-	76.668

Source: Own calculations on EU-SILC 2007 and EU-SILC 2012 dataset.

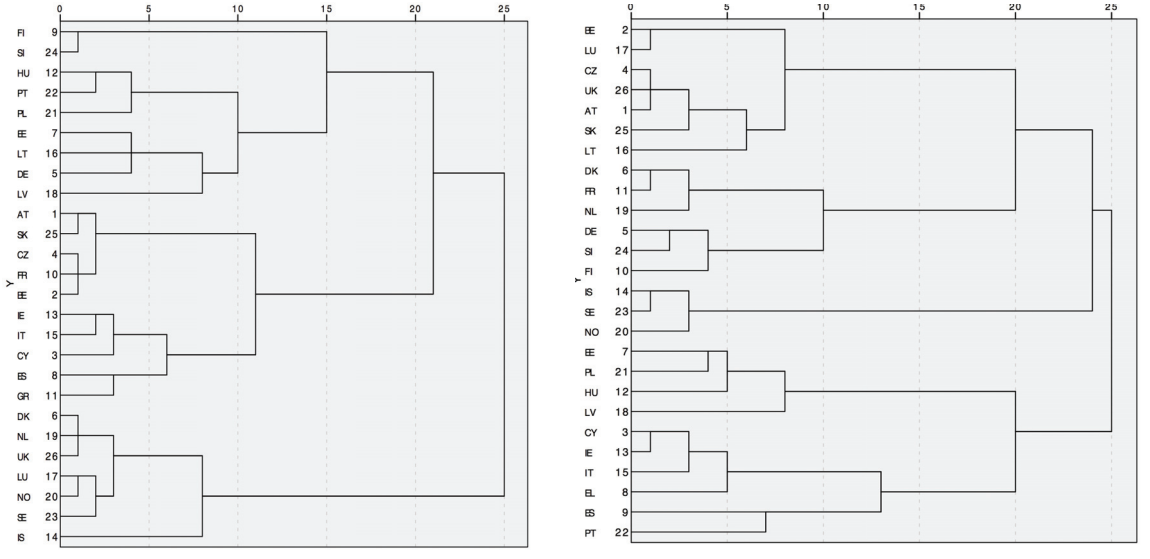
In 2007, nearly 73% of the total variance was explained by synthetizing all our information into three components, among which the first one alone contributes to almost one-third of it all. In 2012, a fourth component needed to be extracted to reach over 76% of the total variance explained, while none of them could yield such a high explanatory power. This could be a hint that the crisis has deeply changed the social environment.

Table A1 in Appendix A shows the structure of the components extracted by the PCA. In the year 2007, the first year of our inquiry, i.e., before the Great Recession, the 11 indicators, in turn yielding evidence for our dimensions of well-being, were reduced into the three components, whose explanatory power was presented in Table 4. The first component contains five indicators (items) related to miserable conditions: 1. Severe housing deprivation, 2. Unmet medical treatment, 3. Unmet dental treatment, 4. Poverty risk, and 5. Extreme material deprivation. The second component contains four indicators all related to the worsening of labor market conditions: 1. Unemployment, 2. Under-employment, 3. Temporary jobs, and 4. NEETs. The third component comprises two indicators regarding health conditions: 1. Suffering from chronic illness, and 2. Activities limited by bad health.

In the year 2012, the Great Recession struck the well-being conditions in many countries and the picture changes accordingly. The first component mainly contains indicators signaling difficulties in the labor market (unemployment, under-employment, and temporary jobs) adding to poverty risk (the only indicator present in the former first component), for a total of four indicators. The second component collects three variables only: NEETs, severe housing deprivation, and extreme material deprivation. This reshuffling of items between the first and second component suggests that the consequences of the financial crisis spread over the weak functioning of the labor market in a multidimensional manner ranging from housing and material deprivation to the NEETs phenomenon. The third component contains two indicators previously included in the first component: unmet medical and dental treatments, while the fourth component, needed to reach a similar level of explained total variance as in 2007, again refers to health problems and is composed by the same indicators included in the third component in 2007. This shows that the impact of the global financial crisis took only a few years to cause a very profound and asymmetric change. How each country contributed to the components is shown in Table A2 in Appendix A, with a graphical representation of the impact of the first and second component (Figure A1).

A cluster analysis delivers a clearer understanding of the PCA results. A Hierarchical Clustering on Principal Components was executed according to Ward's method with Euclidean distance matrix and Z-scores.

The dendrogram plots, clustering the 26 European countries in 2007 and in 2012, are shown in Figure 1a and Figure 1b, respectively. Various groups and subgroups of countries can be identified according to the chosen level of aggregation: the lower the level, the higher the number of sub-groups of countries.



(**a**) 2007 (**b**) 2012

Figure 1. Dendrogram clustering the 26 countries in 2007 (**a**) and 2012 (**b**).

In 2007, the group composed by Denmark, the Netherlands, the United Kingdom, Luxembourg, Norway, Sweden, and Iceland (i.e., mainly Scandinavian and rich countries) stands alone even below the level of aggregation 10. The second main group can be subdivided into two subgroups containing: (1) Germany and many north-eastern countries (Finland, Slovenia, Hungary, Portugal, Poland, Estonia, Lithuania, and Latvia) and (2) France and many central and southern countries (Austria, Slovakia, Czechia, Belgium, Ireland, Italy, Cyprus, Spain, and Greece).

In 2012, a deep reshuffling also occurred among the clusters, so that even the composition of the former two main groups changed. All dimensions, not just the various sources of household material resources, contribute to a differing extent in magnifying the heterogeneity of well-being across Europe.

At a comparable level of aggregation in Figure 1b five subgroups can be identified from the top to the bottom: (1) Belgium, Luxembourg, Czechia, the United Kingdom, Austria, Slovakia, and Lithuania, (2) Denmark, France, the Netherlands, Germany, Slovenia and Finland, (3) Iceland, Sweden, and Norway, all belonging to the first group, (4) Estonia, Poland, Hungary, and Latvia, and (5) Cyprus, Ireland, Italy, Greece, Spain and, Portugal, belonging to the second group.

While the dendrograms gather the 26 countries according to the hierarchical clustering on all principal components, Table A2 in Appendix A provides a breakdown of how each country contributes to the PCA results. The interested reader may check how individual countries fare and compare before and after the Great Recession.

5. Discussion

In the appraisal of our results, one should keep in mind that, going from 2007 to 2012, the item content of the components has changed. In 2007, the first component mainly shows people's conditions for risk of poverty, material deprivation in terms of private goods for a decent everyday life, and the quality of housing conditions as well as some health issues. The second component indicates shortcomings in economic activity related to the labor market. One could think of poor households, which have deprivation as the first problem in their well-being assessment and for whom changes in the labor market and organization would perhaps be able to alleviate their deprived circumstances. In 2012 these issues,

except the NEETs, are included in the first component together with poverty risk, while the second component includes the NEETs with housing and material deprivations. The priority seems to have changed in the consideration of the respondents well-being. It seems that they feel being at risk of poverty for the threat of losing their job. The Great Recession seems to have changed the priorities in the ranking of the difficulties encountered by European households. The contours of disadvantage in 2007, before the crisis, mainly refer to deprivation, which might be caused by difficulties on the labor market, while in 2012 the main weakness is found in troubles in the labor market, leading to poverty. This change does not seem to be short-lived. The analysis of social insecurities by Eurofound (2018), which is based on data gathered in the European Quality of Life Survey (EQLS) across the 28 EU Member States, underlines that, due to the consequences of the 2008 financial crisis, many more people feel insecure about their future. The exposure to vulnerability has reached people who have suddenly become at risk of job loss, or are loaded with over-indebtedness, or are unable to pay for healthcare. Similar results are shown by Ayllón and Gábos (2017) who employ EU-SILC and observe that the three indicators put forward in the Europe 2020 Agenda—being at risk of poverty, severe material deprivation, and low work intensity—present state dependence: in the majority of European countries, once one household is caught by these economic hardships, the persistence is such that is very difficult to escape those conditions, and of course this adds to insecurity. In 2017 at the informal summit held in Gothenburg, the European Pillar of social rights was proclaimed, also envisaging a multidimensional indicator named "at risk of poverty or social exclusion" (AROPE), which includes the sum of people who are either at risk of poverty, or severely materially and socially deprived, or living in a household with a very low work intensity (Atkinson et al. 2017). Some issues characterizing the multidimensional nature of the socio-economic environment have been collected in a single multidimensional indicator summarizing the EU2020 target. Unfortunately the statistics about the single variables have been discontinued, which makes it difficult to extend a full comparison to nowadays.

One of our main findings is that in 2012 employment-related issues stand out as a major determinant of the worsening of the quality of life. This result can be traced back to the negative impact of the demise of EPL of the hoarding of low-skilled workers during the crisis, as argued in Fedotenkov et al. (2024) who employ microdata and specifically address the effects of the Great Recession on labor productivity, detailed by country and sector.

The discussion continues by looking at the different groups of countries.

In trying to make a complex matter simpler, we will keep identifying the groups on their geo-economic position, i.e., Southern, Eastern, Northern, and Western. Exceptions may be checked by looking at the single country's results.

The financial crisis was counteracted with greater success by the richer Western continental countries. The socio-economic performance of the Netherlands, Denmark, Luxembourg, Germany, and Austria was fairly satisfying in 2007 and it looks as if these countries totally recovered by 2012. France and Belgium even more than recovered. The Nordic countries (Iceland, Norway, and Sweden) more than recovered, too.

Eastern European countries had very much benefitted from the economic integration with the German productive system after the demise of the Soviet Union and, to a larger extent, after the 2004 adhesion to the European Union. The recession particularly hit these countries mainly because of a very distressed labor market exacerbating socio-economic conditions; however, deprived people do not show problems in the labor market (Croci Angelini et al. 2020).

In 2007, before the crisis, Hungary and Poland were found in the "bad" quadrant (see Figure A1), the one with difficulties in both first and second components; Slovenia, Slovakia and Czechia had failures in the labor market (second component), while the Baltic countries (Estonia, Latvia, and Lithuania) had no major problems with deprivations (first component).

In 2012, Poland slightly improves, but still remains in the "bad" quadrant, while Hungary joins the Baltic countries. It seems that Slovenia, Slovakia, and Czechia have

reacted to the crisis by struggling to keep the labor market afloat while reducing both deprivations and so reach a position near to zero at the crossing of the two components.

The poor performance of Southern European countries stands out, as the crisis struck them worse of all. Italy and Greece are in the "bad" quadrant in 2007 and are still there in 2012, joined by Cyprus, and aggravate their conditions further away from the zero. Spain and Portugal show improvements in the labor market at the expense of more deprivations caused by deregulation.

In these countries the crisis is far from over also in the following years. Although it has not been possible to compute a recent comparable PCA for the lack of some variables, Southern countries need at least three more years (Portugal) to recover material deprivation levels, while Greece so far has not yet recovered. As for the labor market-related component, unemployment rates in Spain and in Greece were still over 10% in 2018. No other country shows the same record.

Our results are also largely coherent with the findings of Ivanová et al. (2022) who explore EU quality of life through 19 EU-SILC based variables and employ PCA to reduced them to five factors (material-economic conditions, social contacts and existential issues, environmental issues and quality of environment, health limitations, and crime), which are the most important factors affecting EU inhabitants' quality of life. Their study shows the maximum positive correlation (0.93) between households making ends meet with great difficulty and arrears (mortgage or rent, utility bills, or hire purchase); also, the quality of life has been found to be substantially negatively influenced by social insecurity mainly as an effect of economic safety.

The research work by Mazurek (2016) carried out on quarterly macrodata and interested in the magnitude and shape of the Great Recession on 25 EU countries confirms our findings that the periphery of Europe (i.e., mainly East and South) has been the area mostly hit by the Great Recession.

Finally, as for the remaining health-related components, it is perhaps worth mentioning that the performance of none of the Southern countries was in danger in 2007 (together with the Nordic and some other countries), while it was in Portugal in 2012.

All in all, our analysis suggests that the countries with both most valuable economic structure and socio-economic indicators—again the Western continental and the Nordic countries—have shown a remarkable resilience during the period following the financial crisis, confirming their performance both for component 1 and 2. A significant exception is the UK, which is singled out for the worsening of component 2.

Indeed, as shown in Betti (2016) the most relevant change happens in each single group for quality of life rather than in the overall index.

6. Conclusions

The above discussion on the evolution of well-being in Europe indicates that the Great Recession has consolidated the division across the four groups of countries. On the one hand we see that the Central-Western and Nordic countries were able to react to the worsening socio-economic conditions; on the other hand the Central-Eastern countries' convergence has stopped and the Southern countries have further been left behind. The UK is a case in point as its performance has worsened after Brexit in 2016. In other words, the crisis has negatively impinged on the capacity of national and supranational institutions in sustaining the market integration process. The evolution of the well-being conditions has been limited to a strengthening of market integration within the four groups of countries. Due to the impact of the crisis the objective of socio-economic convergence has been set aside.

The purpose of this paper was to evaluate variations in aggregate economic welfare within 26 European countries, by connecting the impact on the main dimensions of well-being of the Great Recession, which was very heterogeneous across Core, Peripheral, and Central-Eastern Europe in particular at the bottom of the income distribution.

In some dimensions, such as risk of poverty, unemployment, and material deprivation, where a recession typically provokes negative effects in the short run, evidence shows that the crisis has worsened the well-being of a sizable number of households, mainly in Southern and Central-Eastern Europe. In some other dimensions, such as health conditions, educational achievements, and housing, the impact of the crisis is not particularly relevant. In the European Union, these indicators regard mainly publicly provided services, which are less subject to the decay that has hit automatic stabilizers after the negative fiscal impulses imposed by Brussels to repair distressed public finances. Typically, the possibly impact of the Great Recession on these dimensions could be assessed only in the medium term, in case the size of the cut to public expenditures would be so large as to gravely worsen the provision of these essential merit goods.

The results of our paper can be compared with those reached by studies aimed at the evaluation of the effects of the Great Recession on European countries (e.g., Mazurek 2016; Fedotenkov et al. 2024) and of the countries' performance towards the EU2020 targets (e.g., Grimaccia 2021). Overall, our results are compatible with them, although the aims, scope, and methods may differ.

As for the method, PCA was applied by Ivanová et al. (2022) over 19 variables to assess the quality of life in member countries of the European Union in eight dimensions reduced to the five most important factors. The advantage of this method lies in the way that a wide amount of information is summarized in few major components which may be more important for policy advice rather than a single figure that is only able to rank the countries.

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

Funding: This research received no external funding.

Data Availability Statement: All EU-SILC data employed in this paper are available at https://ec.europa.eu/eurostat/web/income-and-living-conditions/information-data (accessed on 6 May 2024).

Acknowledgments: While taking joint responsibility for this paper the authors gratefully acknowledge the support, discussion, and valuable comments by Fabio Clementi, Cristina Davino, and Enzo Valentini. We also thank the anonymous referees.

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

Appendix A

In 2007, the 11 indicators listed in the first column of Table A1 were reduced into the three components whose explanatory power is presented in Table 4.

Table A1. Rotated component matrices [a] for the years 2007 [b] and 2012 [c].

	1st Component		2nd Component		3rd Component		4th Component	
	2007	**2012**	**2007**	**2012**	**2007**	**2012**	**2007**	**2012**
1. Severe housing deprivation	**0.750**	0.095	0.269	**0.772**	0.324	0.376	-	0.173
2. Suffer from chronic illness	0.021	−0.224	−0.144	−0.102	**0.870**	0.264	-	**0.835**
3. Unmet medical treatment	**0.900**	0.126	0.011	0.320	0.177	**0.826**	-	0.133
4. Unmet dental treatment	**0.896**	0.241	−0.196	0.087	−0.033	**0.899**	-	−0.101
5. Activities limited by bad health	0.232	0.028	0.186	0.181	**0.834**	−0.231	-	**0.895**

Table A1. *Cont.*

	1st Component		2nd Component		3rd Component		4th Component	
	2007	**2012**	**2007**	**2012**	**2007**	**2012**	**2007**	**2012**
6. Poverty risk	**0.555**	**0.749**	0.291	0.376	−0.076	0.055	-	−0.076
7. Extreme material deprivation	**0.825**	0.208	0.334	**0.891**	0.190	0.275	-	−0.001
8. Unemployment	−0.051	**0.883**	**0.855**	0.269	0.368	0.122	-	−0.062
9. Underemployment	0.372	**0.691**	**0.642**	0.236	0.027	0.428	-	−0.177
10. Temporary jobs	−0.092	**0.730**	**0.735**	−0.395	−0.297	0.126	-	0.004
11. NEETs	0.344	0.071	**0.792**	**0.535**	0.034	−0.001	-	−0.014

Source: Own calculations on EU-SILC 2007 and EU-SILC 2012 dataset. [a] Numbers in bold identify the items composing the different components after the convergence by Varimax method of rotation and Kaiser normalization. [b] rotation reached convergence criteria in 4 iterations. [c] Rotation reached convergence criteria in six iterations.

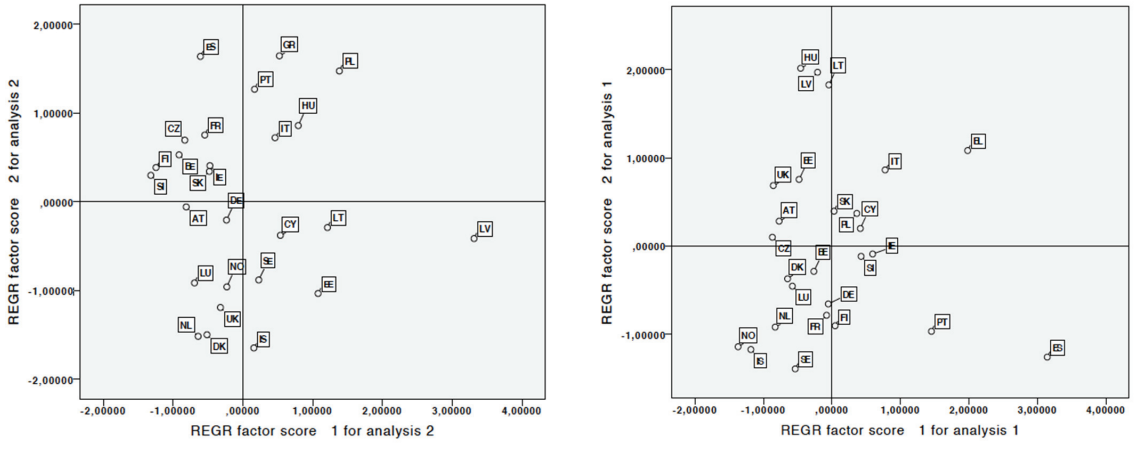
Table A2 shows how the 26 countries contribute to the three (2007) or four (2012) components.

Table A2. Countries' contributions to the principal components.

Country		2007 1st Component	2012 1st Component	2007 2nd Component	2012 2nd Component	2007 3rd Component	2012 3rd Component	2012 4th Component
AT	Austria	−0.8119	−0.7708	−0.05874	0.2812	0.04249	−1.20.05	0.11156
BE	Belgium	−0.91417	−0.26396	0.52803	−0.28846	−0.37991	−0.87335	−1.05017
CY	Cyprus	0.53777	0.41449	−0.37921	0.19816	−1.00462	0.18321	−0.74536
CZ	Czechia	−0.83171	−0.86943	0.69367	0.09863	−0.01061	−0.52352	−0.43346
DE	Germany	−0.2344	−0.05314	−0.20694	−0.65822	1.24639	−0.74793	1.45954
DK	Denmark	−0.51385	−0.6477	−1.49942	−0.37235	−0.30065	−0.14205	0.08512
EE	Estonia	1.07845	−0.47987	−1.03448	0.7549	1.5694	0.32549	1.36469
EL	Greece	0.52409	1.98103	1.6437	1.08254	−1.72526	−0.04791	−0.82843
ES	Spain	−0.60796	3.14244	1.63703	−1.26.12	−1.08996	−0.27919	−0.43654
FI	Finland	−1.24456	0.04651	0.38496	−0.90573	1.95109	0.02104	2.58969
FR	France	−0.5454	−0.07848	0.75137	−0.78641	0.07443	0.39423	0.25036
HU	Hungary	0.79412	−0.45669	0.85781	2.01621	0.91443	0.58656	0.18496
IE	Ireland	−0.47203	0.5942	0.40587	−0.0914	−0.94683	−0.33155	−1.40245
IS	Iceland	0.15807	−1.18332	−1.64814	−1.17504	−2.09799	1.47602	−0.74238
IT	Italy	0.46134	0.77587	0.72078	0.86241	−0.92076	−0.14621	−0.3447
LT	Lithuania	1.21276	−0.0457	−0.29146	1.82744	0.36067	−1.18555	−0.40005
LU	Luxemburg	−0.6931	−0.5774	−0.91574	−0.45614	−0.45531	−0.83776	−1.5566
LV	Latvia	3.31357	−0.2101	−0.41543	1.97079	0.7569	2.7584	0.5087
NL	Netherlands	−0.64051	−0.82843	−1.51746	−0.9206	0.0243	−0.78992	0.82626
NO	Norway	−0.2299	−1.36951	−0.96034	−1.1438	−0.54461	0.52326	−1.20646
PL	Poland	1.38332	0.36321	1.47278	0.36934	−0.0028	1.02713	0.33069
PT	Portugal	0.16762	1.45223	1.26733	−0.96893	0.53069	1.33337	0.26141
SE	Sweden	0.22784	−0.53582	−0.88212	−1.39502	−0.0403	1.49465	−0.41712
SI	Slovenia	−1.31893	0.4255	0.29688	−0.11817	1.7861	−1.32586	1.71076
SK	Slovakia	−0.47952	0.03088	0.34155	0.39451	0.30933	−0.70355	0.4685
UK	United Kingdom	−0.32102	−0.85603	−1.19227	0.68525	−0.04663	−0.98394	−0.58852

Source: Own calculations on EU-SILC 2007 and EU-SILC 2012 dataset.

In Figure A1 the first and second component of both years are plotted in the scatter diagrams 2007 (a) and 2012 (b) into four quadrants. Countries that have the heaviest problems are represented in the upper right quadrant, while countries that fare better are in the double negative lower left quadrant. The two diagrams compare the two years with a warning: in 2007, the first and second components together explain more than half the variance (56%), while in 2012 they only reach 43,5%. A three-dimensional diagram for 2012 would reach 62% (and nearly 73% in 2007) but would not be as reader-friendly.



(**a**) 2007 (**b**) 2012

Figure A1. Scatterplot of the first and second component for the 26 countries, year 2007 (**a**) and 2012 (**b**).

The first component is measured on the horizontal axis and the second component on the vertical axis. Countries represented on the left quadrants show negative values for the first components (i.e., they tend not to be deprived in 2007, nor have major problems on the labor market in 2012), while countries in the lower quadrants show negative values for the second component. Therefore, countries having both negative values tend to fare better than all the rest, while countries having both positive values tend to be more miserable than the remaining countries.

References

Atkinson, Anthony Barnes, and Eric Marlier, eds. 2010. *Income and Living Conditions in Europe*. Luxembourg: Publications Office of the European Union.

Atkinson, Anthony Barnes, Anne-Catherine Guio, and Eric Marlier, eds. 2017. *Monitoring Social Inclusion in Europe*; Luxembourg: Eurostat.

Atkinson, Anthony Barnes, Bea Cantillon, Eric Marlier, and Brian Nolan. 2002. *Social Indicators: The EU and Social Inclusion*. Oxford: Oxford University Press.

Atkinson, Anthony Barnes Eric Marlier, Fabienne Monatigne, and Anne Reinstadler. 2010. Income poverty and income inequality. In *Income and Living Conditions in Europe*. Edited by Anthony Barnes Atkinson and Eric Marlier. Luxembourg: Publications Office of the European Union.

Ayllón, Sara, and Andras Gábos. 2017. The Interrelationships between the Europe 2020 Poverty and Social Exclusion Indicators. *Social Indicators Research: An International and Interdisciplinary Journal for Quality-of-Life Measurement* 130: 1025–49. [CrossRef]

Betti, Gianni. 2016. What impact has the economic crisis had on quality of life in Europe? A multidimensional fuzzy approach. *Quality and Quantity* 51: 351–64. [CrossRef]

Crimmins, Eileen, Jung Kj Kim, and Teresa Seeman. 2009. Poverty and Biological Risk: The Earlier "Aging" of the Poor. *Journal of Gerontology: Medical Sciences* 64A: 286–92. [CrossRef] [PubMed]

Croci Angelini, Elisabetta. 2015. Titolo di godimento della casa ed eterogeneità delle condizioni economiche e abitative nei paesi europei. In *La casa. il benessere e le diseguaglianze*. Edited by Francesco Farina and Maurizio Franzini. Milano: Egea.

Croci Angelini, Elisabetta, and Alessandra Michelangeli. 2012. Axiomatic measurement of multidimensional well-being inequality: Some distributional questions. *The Journal of Socio-Economics* 41: 548–57. [CrossRef]

Croci Angelini, Elisabetta, Francesco Farina, and Enzo Valentini. 2016. Contagion across Eurozone's sovereign spreads and the Core-Periphery divide. *Empirica* 43: 197–213. [CrossRef]

Croci Angelini, Elisabetta, Francesco Farina, and Enzo Valentini. 2020. Wage and employment by skill levels in technological evolution of South and East Europe. *Journal of Evolutionary Economics* 30: 1497–514. [CrossRef]

Decancq, Koen, and Maria Ana Lugo. 2013. Weights in Multidimensional Indices of Wellbeing: An Overview. *Econometric Reviews* 32: 7–34. [CrossRef]

De Grauwe, Paul, and Yuemei Ji. 2013. Self-fulfilling crises in the Eurozone: An empirical test. *Journal of International Money and Finance* 34: 15–36. [CrossRef]
EU Council of Ministers. 1985. *Council Decision of 19 December 1984 on Specific Community Action to Combat Poverty (85/8/EEC)*. OJEC. L2. Brussels: EU Council of Ministers.
Eurofound. 2018. *Social Insecurities and Resilience*. Luxembourg: Publications Office of the European Union.
European Commission. 2010. *Europe 2020. A Strategy for Smart, Sustainable and Inclusive Growth*. Brussels: COM(2010) 2020 final. Available online: https://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=COM:2010:2020:FIN:en:PDF (accessed on 6 May 2024).
Fedotenkov, Igor, Virmantas Kvedaras, and Miguel Sanchez-Martinez. 2024. Employment protection and labour productivity growth in the EU: Skill-specific effects during and after the Great Recession. *Empirica* 51: 209–62. [CrossRef]
Figari, Francesco, Andrea Salvatori, and Holly Sutherland. 2010. Policy simulation across countries using EUROMOD: Stress testing European welfare systems for unemployment. In *Income and Living Conditions in Europe*. Edited by Anthony Barnes Atkinson and Eric Marlier. Luxembourg: Publications Office of the European Union.
Freudenberg, Michael. 2003. *Composite Indicators of Country Performance: A Critical Assessment*. OECD Science, Technology and Industry Working Papers, No. 2003/16. Paris: OECD Publishing. [CrossRef]
Fusco, A., Anne-Catherine Guio, and Eric Marlier. 2010. Characterising the income poor and the materially deprived in European countries. In *Income and Living Conditions in Europe*. Edited by Anthony Barnes Atkinson and Eric Marlier. Luxembourg: Publications Office of the European Union.
Grimaccia, Elena. 2021. Europe 2020 Strategy for a Smart, Inclusive and Sustainable Growth: A First Evaluation. *Rivista Italiana di Economia Demografia e Statistica* 75: 65–77.
Hernández-Quevedo, Cristina, Cristina Masseria, and Elias Mossialos. 2010. Socio-economic determinants of health in Europe. In *Income and Living Conditions in Europe*. Edited by Anthony Barnes Atkinson and Eric Marlier. Luxembourg: Publications Office of the European Union.
Horemans, Jeroen, Ive Marx, and Brian Nolan. 2015. *Hanging in, but Only Just. Part-Time Employment and in-Work Poverty throughout the Crisis*. CSB Working Paper n.3. Antwerp: Herman Deleeck Centre for Social Policy, University of Antwerp.
ILO. 2015. *Wage and Income Inequality*. Global Wage Report 2014/15. Geneva: ILO.
Ivanová, Monika, Radoslav Klamár, and Erika Feckova Škrabuľáková. 2022. Identification of Factors Influencing the Quality of Life in European Union Countries Evaluated by Principal Component Analysis. *Geographica Pannonica* 26: 13–29. [CrossRef]
Jenkins, Stephen P., Andrea Brandolini, John Micklewright, and Brian Nolan. 2013. *The Great Recession and the Distribution of Household Income*. Oxford: Oxford University Press.
Kemeny, Jim. 2001. Comparative housing and welfare: Theorising the relationship. *Journal of Housing and the Built Environment* 16: 53–70. [CrossRef]
Mazurek, Jiri. 2016. The Evaluation of Recession Magnitudes in EU Countries during the Great Recession 2008–2010. *Review of Economic Perspectives* 16: 231–44. [CrossRef]
Murtin, Fabrice, and Marco Mira d'Ercole. 2015. *Household Wealth Inequality across OECD Countries: New OECD Evidence*. OECD Statistical Brief. n.21. Paris: OECD.
Nardo, Michela, Michaela Saisana, Andrea Saltelli, Stefano Tarantola, Anders Hoffmann, and Enrico Giovannini. 2008. *Handbook on Constructing Composite Indicators: Methodology and User Guide*. OECD JRC47008. Paris: OECD.
Nolan, B., and Christopher T. Whelan. 2010. Using Non-Monetary Deprivation Indicators to Analyse Poverty and Social Exclusion: Lessons from Europe? *Journal of Policy Analysis and Management* 29: 305–25. [CrossRef]
Nolan, Brian, and Christopher T. Whelan. 2014. The Social Impact of Income Inequality: Poverty, Deprivation and Social Cohesion. In *Changing Inequalities in Rich Countries: Analytical and Comparative Perspectives*. Edited by Wiemer Salverda, Brian Nolan, Daniele Checchi, Ive Marx, Abigail McKnight, Istaván György Tóth and Herman van de Werfhorst. Oxford: Oxford University Press.
Nussbaum, Martha, and Amartya Kumar Sen, eds. 1993. *The Quality of Life*. Oxford: Oxford University Press.
OECD. 2015. *Employment Outlook*. Paris: OECD.
Sauli, Hannele, and Veli-Matti Törmälehto. 2010. The distributional impact of imputed rent. In *Income and Living Conditions in Europe*. Edited by Anthony B. Atkinson and Eric Marlier. Luxembourg: Publications Office of the European Union.
Sen, Amartya Kumar. 1985. *Commodities and Capabilities*. Amsterdam and Oxford: North-Holland.
Sierminska, Eva. 2012. *Wealth in the Crisis*. Social Situation Observatory. Income Distribution and Living Conditions. Research Note 9. Brussels: European Commission.
Sologon, Denisa Maria, and Cathal O'Donoghue. 2014. Shaping Earnings Insecurity: Labour Market Policy and Institutional Factors. *Review of Income and Wealth* 60: S205–32. [CrossRef]
Stiglitz, Joseph E., Amartya Kumar Sen, and Jean-Paul Fitoussi. 2009. *Report by the Commission on the Measurement of Economic Performance and Social Progress*. New York: Columbia University.
Weymark, John A. 2006. The normative approach to the measurement of multidimensional inequality. In *Inequality and Economic Integration*. Edited by Francesco Farina and Ernesto Savaglio. London: Routledge, pp. 303–28.

Whelan, Christopher T., and Bertrand Maître. 2012. Understanding Material Deprivation in Europe: A Multilevel Analysis. *Research in Social Stratification and Mobility* 30: 489–503. [CrossRef]
Whelan, Christopher T., Brian Nolan, and Bertrand Maitre. 2014. Multidimensional Poverty Measurement in Europe: An Application of the Adjusted Headcount Approach. *Journal of European Social Policy* 24: 183–97. [CrossRef]

economies

MDPI

Article

The Fabric of Transition: Unraveling the Weave of Labor Dynamics, Economic Structures, and Innovation on Income Disparities in Central and Eastern Europe Nations

Adriana AnaMaria Davidescu [1,2,*], Oana-Ramona Lobonţ [3] and Tamara Maria Nae [4,5]

[1] Department of Statistics and Econometrics, The Faculty of Economic Cybernetics, Statistics and Informatics, The Bucharest University of Economic Studies, 010552 Bucharest, Romania
[2] Department of Education, Training and Labour Market, National Scientific Research Institute for Labour and Social Protection, 010643 Bucharest, Romania
[3] Department of Finance, Faculty of Economics and Business Administration, West University of Timisoara, 300223 Timisoara, Romania; oana.lobont@e-uvt.ro
[4] Department of Economics and Economic Policies, Faculty of Theoretical and Applied Economics, Bucharest University of Economic Studies, 010552 Bucharest, Romania; tamara.nae@economie.ase.ro
[5] Research and Analysis Department, Ministry of Finance, 050706 Bucharest, Romania
* Correspondence: adriana.alexandru@csie.ase.ro

Abstract: In recent years, the issue of income inequality has ascended to the forefront of national and international agendas, underscored by the urgency to navigate the complexities of market-driven economies without exacerbating social disparities. These challenges are particularly pronounced in the post-communist nations of Central and Eastern Europe, where the transition legacy and the marketization forces present unique dynamics in the evolution of income disparities. This research investigates the intricate mechanisms through which marketization impacts income inequality within the Central and Eastern European countries context, aiming to uncover how economic transformations influenced by global sustainability goals can contribute to narrowing the income gap. By employing panel data estimation techniques and Generalized Method of Moments (GMM) analysis, this study highlights the enduring nature of income disparities and the critical roles played by economic growth, education investment, labor market reforms, globalization, and governance quality in shaping equitable income distributions. Findings reveal that, despite the competitive nature of market economies potentially creating disparities, strategic policy interventions in education, economic policy, and labor market regulations can mitigate the adverse effects of marketization on income inequality. Additionally, this research emphasizes the importance of strong institutional frameworks and the nuanced role of the informal economy in influencing income distribution dynamics.

Keywords: income inequality; determinants; CEE countries; panel data approach; GMM; social progress; convergence

Citation: Davidescu, Adriana AnaMaria, Oana-Ramona Lobonţ, and Tamara Maria Nae. 2024. The Fabric of Transition: Unraveling the Weave of Labor Dynamics, Economic Structures, and Innovation on Income Disparities in Central and Eastern Europe Nations. *Economies* 12: 68. https://doi.org/10.3390/economies12030068

Academic Editor: Fabio Clementi

Received: 12 December 2023
Revised: 28 February 2024
Accepted: 4 March 2024
Published: 14 March 2024

1. Introduction

Inequality has been the subject of a great debate at all times. This phenomenon has gained particular interest among economists since the economic downturns that hit Europe after the most significant wave of accession. Concurrently, there has been a heightened emphasis in research on income distribution, which is currently emerging as a progressively urgent economic and social concern. This is especially notable in emerging European nations, where inequalities surpass the mean observed in the European Union (EU).

In the aftermath of the collapse of communist regimes in Central and Eastern Europe, the interplay between income inequality and marketization has witnessed significant changes. The shift from centrally planned to market-driven economies in the region has presented both opportunities and challenges. While marketization has stimulated economic growth, fostering an overall increase in prosperity and the emergence of a

growing middle class, the swift implementation of market reforms has concurrently given rise to varying levels of income inequality. Certain segments of the population have been more adept at leveraging new economic opportunities, leading to disparities that need careful consideration.

Reducing income inequality in Central and Eastern European countries holds paramount significance for fostering sustainable development, social cohesion and resilience. High levels of income inequality can undermine a region's economic potential by limiting access to education, healthcare, and opportunities for a significant portion of the population. Addressing inequality contributes to political stability and a more inclusive society. Moreover, a more equitable distribution of income can stimulate domestic demand, fostering a robust and resilient economy. As CEE nations continue to navigate the challenges of transition and marketization, prioritizing policies that reduce income disparities becomes crucial for building a prosperous and harmonious future for their citizens.

The genesis of this study is rooted in the persistent challenges pertaining to inequality, notably discernible within the European Union and accentuated within CEE countries. The research is motivated by the identification of a comprehensive set of policy recommendations designed to ameliorate the social landscape in CEE nations, with a strategic aim of mitigating and ultimately eradicating the socio-economic disparities between Eastern Europe and the rest of European Union.

This paper contributes significant added value to the existing literature by offering a nuanced examination of the complex interplay between marketization, income inequality, and institutional quality, with a specific focus on post-communist countries in CEE. The study not only investigates the impact of marketization on income distribution but also integrates the crucial dimension of institutional quality, providing a more comprehensive understanding of the factors influencing inequality dynamics in the CEE region. By identifying key policy recommendations tailored to the unique socio-economic landscape of CEE countries, this paper offers actionable insights for policymakers striving to address and mitigate income disparities.

This paper adopts a systematic structure, commencing with an insightful introduction that articulates the research problem's significance and delineates the study's objectives. Subsequently, this literature review meticulously examines existing scholarship, establishing a robust theoretical framework. The Data and Methodology section outlines the research design, data sources, and analytical approach, encompassing the nuanced exploration of random effects, fixed effects, and rigorous endogeneity testing techniques. Moving forward, the Results and Discussion section synthesizes empirical findings, differentiating between random and fixed effects. The conclusion succinctly summarizes key findings, underscores their contributions to the existing body of knowledge, and propounds avenues for future research. Additionally, the paper culminates with judicious policy recommendations, deriving practical implications from the study's insights and offering guidance for decision makers in relevant domains. This cohesive structure ensures a logical progression of ideas, facilitating a comprehensive and impactful presentation of the research.

2. Literature Review

2.1. Marketization and Income Inequality: A Complex Nexus

The relationship between marketization and income inequality is highly contextual, showing variance across different countries and regions. This variance is largely shaped by specific policy measures, institutional frameworks, and socio-economic conditions unique to each context. Scholars have underscored the critical importance of considering the distributive impacts of market-oriented policies, which include changes in access to education, social protection, and employment opportunities. The consensus from these studies suggests an intricate and multifaceted link between income inequality and marketization, highlighting the necessity for a nuanced understanding that accommodates a broad spectrum of contextual factors.

2.2. Labor Market Dynamics and Their Disparate Impacts

The labor market plays a pivotal role in influencing income inequality, with technological advancements and shifts in the demand for skilled labor contributing significantly to wage disparities. Acemoglu and Autor (2011) and Piketty (2014) provide evidence of how economic growth periods and technological shifts exacerbate income inequality. Goldin and Katz (2007) further this discussion by emphasizing the exacerbating role of education and skill differentials within the labor market. The importance of labor market policies, such as minimum wage regulations and social protection measures in shaping income distribution, is highlighted by Atkinson and Morelli (2010), with Chetty et al. (2014) discussing the persistence of inequalities across generations due to labor market opportunities.

Card and Krueger (1994) and Autor et al. (2008) delve into the effects of minimum wage policies and labor market polarization, underscoring the significance of educational composition in the workforce and its impact on income disparities. The distribution of employees across industries, as discussed by Goos and Manning (2007), demonstrates how technological advancements have led to the decline in middle-skilled jobs, further increasing income inequality.

2.3. Economic Performance, Structure, and Income Inequality

The intricate relationship between economic growth and income inequality has captured scholarly attention, with mixed findings regarding its impacts. Barro (2000) and Forbes (2000) explore this relationship, while Berg and Ostry (2011) suggest that extreme income disparities may disrupt economic stability. Persson and Tabellini (1994) underscore the mediating role of institutional quality in this relationship, proposing that well-designed institutions can alleviate the adverse effects of inequality on development.

2.4. Openness of the Economy: A Double-Edged Sword

The interplay between economic openness and income inequality has been extensively studied, with Rodrik (1997) and Milanovic (2005a, 2005b) discussing how trade liberalization and globalization can initially increase income inequality. Bergstrand and Egger (2007), along with Firebaugh and Goesling (2004), emphasize the contingent nature of this relationship on factors like development level and institutional quality.

2.5. Shadow Economy and Income Disparities

The shadow economy's role in influencing income inequality is highlighted by Schneider and Enste (2000) and Torgler and Schneider (2007), noting the informal sector's contribution to wage disparities and the growth of informal employment driven by income inequality. Buehn and Schneider (2012) stress the importance of the institutional context in understanding these dynamics.

2.6. Technological Advancements and High-Tech Exports

The literature on high-tech exports and income inequality presents a nuanced view, with Lin and Li (2011) and Gouvea and Wang (2019) discussing the sector's potential to both exacerbate and mitigate income disparities. The importance of investments in education and technology in narrowing skill differentials is noted by Barro (2000), with Li and Liu (2005) cautioning about the uneven benefits of high-tech exports.

2.7. Governance, Institutional Quality, and Income Distribution

The role of governance and institutional quality in addressing income inequality is emphasized by Acemoglu and Robinson (2012) and Kaufmann et al. (2010), highlighting the importance of strong institutions in promoting equitable resource distribution. Murtin and Wacziarg (2014) provide empirical evidence linking improvements in institutional quality to reductions in income inequality.

Despite extensive research on individual marketization factors and their impact on income inequality, there exists a notable gap in studies that provide a holistic analysis

integrating these elements into a unified framework. The current literature often examines these aspects in isolation, lacking a comprehensive understanding of the synergies and interactions among marketization components and their collective influence on income distribution. There's a critical need for research that bridges these gaps, offering an integrated perspective that encompasses the dynamic interplay among labor market dynamics, economic performance, technological shifts, and institutional frameworks in shaping income inequality outcomes. Addressing this void is imperative for policymakers and scholars seeking a thorough comprehension of marketization's multifaceted impact on income distribution.

The description of the variables that will be used in the empirical analysis can be studied in (Table 1).

Table 1. List of the variables and data source.

Variable	Source	Sign
Endogenous		
Gini Coefficient (pp)	Eurostat data base	−
Exogenous variable		
Economic Performance and Labor market		
Minimum monthly wage (%)	Eurostat data base	−
Strictness of employment protection index, individual and collective dismissals (%)	Employment Protection Database, OECD	−
Gross domestic product per capita (euro/cap.)	Eurostat data base	(+/−)
Economic growth/cap. (%)	Eurostat data base	(+/−)
Employed population with tertiary education (%)	Eurostat data base	−/+
Employees in the industry (%)	Eurostat data base	−
Education expenditure (% GDP)	Eurostat data base	(−)
Informal economy (% GDP)	Global Economy	(−/+)
Globalization		
Share of high-tech exports (%) of Total Exports	Eurostat data base	+/−
Innovation index (%)	Global Economy	(+/−)
Openness of the economy (% GDP)	Eurostat data base	(+/−)
Quality of institutions		
Regulatory quality (pp)	World Bank	(−)
Rule of law (pp)	World Bank	(−)
Control of Corruption (pp)	World Bank	(−/+)

3. Data and Methodology

Addressing social issues and enhancing fair income distribution necessitates a deep dive into the factors influencing income disparity. This exploration is crucial for fostering broader socio-economic inclusion, elevating the general quality of life, and ensuring economic and social stability, which in turn, strengthens socio-economic cohesion and resilience against future crises.

The empirical component of this study zeroes in on the determinants of income inequality within ten CEE countries, excluding Croatia due to data limitations. Utilizing panel data regression analysis, the research covers annual data from 2008 to 2019, dissecting the influence of identified determinants across four main categories: labor market institutions, economic development, globalization, and governance. The model employed is given by:

$$\text{Gini}_{it} = \beta 0 + \beta 1 * \text{Labour Market Institutions}_{it} + \beta 2 * \text{Economic Development}_{it} + \beta 3 * \text{Globalisation}_{it} + \beta 4 * \text{Governance}_{it} + \beta 5 * \text{Control Variables}_{it} + \varepsilon t \quad (1)$$

where i represents the cross sections, t the period, and β the coefficients of influencing factors on income distribution. Based on the literature and empirical evidence, the analysis includes additional control variables to account for factors such as sectoral employment distribution, urbanization effects, and inflation impacts on income inequality.

Table 1 comprehensively explains the variables, their definitions, and the data sources utilized. These sources encompass the Eurostat database. Meanwhile, Table 2 furnishes descriptive details about the main leading indicators.

Table 2. Empirical results of income inequality determinants with random effects models.

Variables	M1	M2	M3	M4	M5	M6
Economic Performance and Labor market						
MMWBI	−0.014 (0.0016) (0.00)	−0.01 (0.001) (0.00)	−0.01 (0.002) (0.00)	0.0005 (0.002) (0.79)	−0.001 (0.00) (0.30)	
MMWBI × URBANISATION	0.0001 (3.86) (0.00)		0.0001 (3.76) (0.01)			
ECG/cap.	−0.30 (0.15) (0.07)	−0.53 (0.20) (0.02)				0.04 (0.03) (0.28)
SHADOW_EC	−0.002 (0.08) (0.97)	0.18 (0.05) (0.00)			0.05 (0.08) (0.54)	
SHADOW_EC × ECG/cap.	0.02 (0.00) (0.02)	0.02 (0.008) (0.01)	0.003 (0.002) (0.10)			
TERED	0.44 (0.03) (0.00)	0.34 (0.02) (0.00)	0.30 (0.02) (0.00)	0.03 (0.06) (0.61)	0.09 (0.06) (0.10)	0.01 (0.04) (0.80)
EMP_IND			−0.53 (0.08) (0.00)	−0.79 (0.12) (0.00)	−0.82 (0.11) (0.00)	−0.85 (0.10) (0.00)
ED_SPEND			−1.15 (0.34) (0.00)	−0.66 (0.35) (0.06)	−0.93 (0.50) (0.09)	−0.75 (0.32) (0.02)
Globalization						
INNOV		−0.16 (0.06) (0.03)		−0.17 (0.07) (0.02)	−0.15 (0.08) 0.09)	−0.14 (0.08) (0.09)
HIGHTECHXP			0.003 (0.04) (0.93)			−12 (0.07) (0.10)
OPENESS	−0.07 (0.00) (0.00)	−0.06 (0.003) (0.00)	−0.05 (0.004) (0.00)			

Table 2. *Cont.*

Variables	M1	M2	M3	M4	M5	M6
Quality of Institutions						
REG_QUAL						
RULE_OF_LAW	−3.43 (0.50) (0.00)	−0.23 (0.70) (0.75)	−1.49 (0.50) (0.01)		−1.09 (0.08) (0.30)	
CONT_CORR				−1.16 (1.16) (0.18)		
Constant	34.15 (1.55) (0.00)	42.64 (3.25) (0.00)	50.43 (2.65) (0.00)	57.16 (5.02) (0.00)	56.52 (2.71) (0.00)	58.64 (4.13) (0.00)
Obs.no.	120	120	120	120	120	120
F-test	211.63 (0.00)	39.53 (0.00)	90.94 (0.00)	11.39 (0.00)	20.39 (0.00)	14.12 (0.00)
S.E. of Reg.	1.177	1.89	1.81	1.48	1.64	1.51
Adj. R^2	0.945	0.86	0.87	0.34	0.53	0.39
Lagrange Multiplier Tests for Random Effects Null hypotheses: No effects						
Breusch–Pagan Multiplier LM test	21.06 (0.00)	19.05 (0.00)	10.93 (0.00)	138.92 (0.00)	68.93 (0.00)	117.12 (0.00)
Testing the normality						
Jarque–Bera	5.09 (0.07)	7.72 (0.02)	3.08 (0.21)	2.74 (0.25)	2.01 (0.36)	2.57 (0.27)
Testing for cross-sectional dependence/contemporaneous correlation: using Breusch–Pagan LM test of independence						
Breusch–Pagan LM	111.54	109.25	95.59 (0.00)	77.72 (0.00)	95.96	76.57
	(0.00)	(0.00)			(0.00)	(0.00)
Pesaran Scaled LM	7.01	6.77	5.33	3.44	5.37	3.32
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Pesaran CD	1.10	0.93	0.92	3.42	3.46	2.97
	(0.27)	0.35	0.35	(0.00)	(0.00)	(0.00)
Testing for heteroskedasticity						
Panel Cross Section Heteroskedasticity L.R. test	57.59 (0.00)	54.24 (0.00)	48.34 (0.00)	81.70 (0.00)	67.26 (0.00)	45.54 (0.00)

Note: Within the table, the coefficients are displayed together with standard errors and the probabilities within (). Standard errors are typically displayed in parentheses right below the coefficients to indicate they are related but distinct values.

The primary constraints of the empirical analysis arise from the lack of data availability for post-communist nations. The most recent estimate by Medina and Schneider (2018) pertains to the year 2015. Consequently, for the period spanning from 2015 to 2019, we relied on this latest estimate. Regarding the innovation index, due to data unavailability for the years 2008 to 2011, we utilized the 2011 value as a substitute for this timeframe. Additionally, data regarding the share of high-tech exports was only accessible up to 2018; consequently, we maintained the same metrics for the year 2019 due to the unavailability of updated information.

To rigorously examine the impact of various factors on income inequality within the context of CEE countries, our empirical analysis employs a multi-faceted econometric

approach that integrates fixed effects (FE), random effects (RE), and Difference Generalized Method of Moments (Dif-GMM) models. This combination allows for a comprehensive understanding of the dynamics at play, offering distinct advantages in addressing specific data and econometric challenges.

The general model for analyzing income inequality, represented by the Gini coefficient (Gini it), across CEE countries over time is specified as:

$$\text{Gini}_{it} = \beta 0 + \beta 1 X_{1it} + \beta 2 X_{2it} + \ldots + \beta k X_{k\,it} + u_i + v_t + \varepsilon_{it} \quad (2)$$

where:

i = index countries
t = index time (years)
$X_{1it} \ldots X_{k\,it}$ are the explanatory variables that include labor market institutions, economic development, globalization, governance, and other control variables.
$\beta 0$ is the intercept
$\beta 1 \ldots \beta k$ are the coefficients of the explanatory variables
u_i is the unobserved country specific effect
v_t is the unobserved time specific effect
ε_{it} is the idiosyncratic error term

In our estimation process, we initially utilized cross-section and period fixed effects models in combination with the ordinary least squares (OLS) technique. These models were progressively fine-tuned to capture income disparity effectively across all CEE nations. To determine the most suitable model, we conducted Redundant Fixed Effects and Hausman tests, evaluating the choice between fixed effects models (FEM) or random effects models (REM). Moreover, we employed the Breusch–Pagan Lagrange (LM) multiplier to evaluate the random effects' consistency and select between a random effects regression and a conventional OLS regression.

Key tests—Redundant Fixed Effects, Hausman, and Breusch–Pagan LM—inform the selection between fixed effects and random effects models, ensuring the model's fit and reliability.

Challenges such as heteroskedasticity and autocorrelation are addressed through specific statistical tests and corrections, including Breusch–Pagan for heteroskedasticity, upholding the integrity of regression outcomes. Cross-sectional dependency is scrutinized via Pesaran's test among others, with heteroskedasticity and normality of residuals assessed to ensure robust statistical inferences.

Recognizing the dynamic nature of inequality and potential endogeneity issues, this study employs advanced econometric techniques. The integration of Dif-GMM with FE and RE models in the analysis of income inequality within CEE countries offers a robust methodological framework that mitigates specific econometric issues inherent in panel data analysis. This approach enhances the reliability of the empirical findings by effectively controlling for unobserved heterogeneity, addressing endogeneity, and ensuring the consistency and efficiency of the estimated coefficients.

The Dif-GMM approach transforms the original model by differencing:

$$\Delta\text{Gini}_{it} = \Delta\beta 1 X_{1it} + \Delta\beta 2\, X_{2it} + \ldots + \Delta\beta k X_{k\,it} + \Delta\varepsilon_{it} \quad (3)$$

where Δ denotes the first difference operator. The Dif-GMM then uses lagged levels of the variables as instruments for the differenced equations, effectively addressing endogeneity.

The FE model controls for unobserved heterogeneity when this heterogeneity is constant over time but varies between entities. It effectively captures the impact of variables that change over time within entities, removing the effect of time-invariant characteristics. The RE model is useful when the unobserved heterogeneity is assumed to be uncorrelated with the explanatory variables. It allows for generalization beyond the sampled entities and is more efficient than the FE model if the unobserved effects are indeed random. The

Dif-GMM estimator, developed by Arellano and Bond, is particularly advantageous in addressing the endogeneity problem that often plagues panel data analyses.

Therefore, this methodology section outlines a comprehensive approach to analyzing income inequality in CEE countries, leveraging panel data regression, addressing methodological challenges, and employing advanced econometric techniques to uncover the multifaceted determinants of income disparity. By meticulously handling data constraints, methodological issues, and endogeneity concerns, this study aims to contribute valuable insights into policy formulation to reduce income inequality and foster socio-economic inclusion.

4. Results and Discussion

4.1. Investigating the Impact of Marketization Factors on Income Distribution in CEE Countries

This section aims to analyze and understand how various marketization processes influence the distribution of income within the economies of Central and Eastern Europe. Integrating insights from the literature review into the empirical findings of the study on the impact of marketization factors on income distribution in CEE countries allows for a nuanced understanding of how these dynamics play out in specific regional contexts. The literature underscores the complexity of the relationship between marketization and income inequality, emphasizing the critical role of policy measures, institutional frameworks, and socio-economic conditions.

This study finds that higher minimum wages are generally associated with reduced income inequality in CEE countries, aligning with Acemoglu and Autor's (2011) and Card and Krueger's (1994) discussions on wage disparities and the redistributive effects of minimum wage policies. The impact of labor market dynamics, as detailed by Goldin and Katz (2007), further supports the empirical evidence, suggesting that policies aimed at reducing skill differentials and protecting low-wage workers are vital for mitigating income inequality. The mixed impact of economic growth on income inequality reflects the literature's varied perspectives, with initial growth potentially reducing inequality, as supported by Barro (2000) and Forbes (2000), before the benefits become unevenly distributed. Berg and Ostry's (2011) suggestion that extreme disparities might hinder economic stability complements the empirical evidence, indicating the importance of balanced growth and inclusive policies.

The negative coefficients for economic openness in reducing income inequality are consistent with the discussions by Rodrik (1997) and Milanovic (2005a) and Milanovic (2005b), highlighting globalization's complex role in shaping income distribution. The findings resonate with Bergstrand and Egger's (2007) emphasis on the contingent nature of globalization's impacts, suggesting that openness can benefit income distribution when coupled with strong institutional frameworks and development strategies.

The study's indication that high-tech exports contribute to reducing income inequality aligns with Lin and Li's (2011) observations on the sector's dual potential to affect income disparities. This underscores the importance of investments in education and technology, as noted by Barro (2000), in leveraging high-tech industries for equitable growth. Also, the significant negative impact of improved governance and anti-corruption measures on income inequality echoes the literature's consensus on the importance of strong institutions, as discussed by Acemoglu and Robinson (2012) and Kaufmann et al. (2010). These findings highlight the critical role of institutional quality in ensuring equitable resource distribution and mitigating the adverse effects of marketization on income inequality.

This integrated analysis contributes to filling the gap in the literature by offering a holistic perspective that combines labor market dynamics, economic performance, globalization effects, and institutional quality into a unified framework for understanding income inequality. It underscores the need for comprehensive policy approaches that consider the interplay among these factors to effectively tackle income disparities in the CEE region and beyond. The main empirical findings of the random effects models are outlined in Table 2.

4.2. Endogeneity Testing of the Core Factors Influencing Income Distribution Dynamics in Emerging Countries

Reinterpreting the empirical findings from the GMM estimation through the lens of the integrated literature insights and empirical data, we derive a nuanced understanding of income distribution dynamics in CEE countries. This section revisits the core determinants of income inequality, emphasizing the role of dynamic panel data estimation in mitigating endogeneity and revealing the temporal influences on income disparities.

The significant impact of lagged income inequality variables across all models not only evidences the enduring nature of income disparities but also suggests a compounding effect over time. This temporal persistence underscores the inherent challenges in effecting swift changes in income distribution patterns.

The variable impact of minimum wage adjustments on income inequality reflects its dual nature within different economic contexts, underscoring the complexity of wage policy outcomes. These findings illuminate the conditionality of minimum wage effectiveness on urban economic structures and labor market dynamics, challenging the traditional view of minimum wage as a straightforward tool for reducing income disparities.

The inverse relationship between economic growth and income inequality underscores the potential of inclusive growth strategies to enhance social welfare. This suggests that economic expansion, when aligned with equitable social policies, can significantly mitigate income disparities by broadening access to economic benefits.

Increased allocations for education emerge as a potent mechanism for reducing income inequality, emphasizing the transformative power of education in leveling socioeconomic disparities. This pivotal finding aligns with the consensus on education as a foundational pillar for equitable development and opportunity access.

The reduction in income inequality associated with economic openness and high-technology exports underscores the potential of globalization to foster a more equitable income distribution. This suggests that strategic integration into the global economy, through high-value sectors, can catalyze inclusive economic growth.

The substantial negative impact of governance improvements and anti-corruption measures on income inequality highlights the indispensable role of robust institutional frameworks in promoting fair economic distribution. This aligns with the principle that effective governance and transparent institutions are fundamental to equitable development.

The nuanced role of the shadow economy, with its potential to both alleviate and exacerbate income disparities, underscores the complex interplay between formal and informal economic sectors. This finding suggests that the informal economy can serve as both a cushion and a challenge in the quest for equitable income distribution, emphasizing the need for policies that recognize and address its multifaceted impacts.

Integrating GMM estimation results with comprehensive literature insights offers a rich understanding of the multifaceted drivers of income inequality in CEE countries. It highlights the importance of a holistic policy approach that encompasses economic growth, labor market reforms, educational investment, global integration, and institutional strengthening. Furthermore, it underscores the need to consider the informal economy's intricate role in shaping income distribution dynamics. This integrated perspective is crucial for formulating policies that not only target economic indicators but also address the underlying structural and institutional determinants of income inequality, paving the way for more inclusive and resilient economic systems. The empirical findings from the GMM estimation, which investigates the primary drivers of income inequality in CEE countries, are outlined in Table 3.

Table 3. Endogeneity testing of the main determinants of income inequality in CEE countries using dynamic panel data estimation.

	M1	**M2**	**M3**	**M4**	**M5**	**M6**
Lagged dep.variable (Gini t − 1)	0.622 (0.103) (0.00)	0.60 (0.10) (0.00)	0.55 (0.11) (0.00)	0.57 (0.08) (0.00)	0.57 (0.09) (0.00)	−0.40 (0.79) (0.61)
Economic Performance and Labor market						
MMWBI	0.0044 (0.004) (0.29)	0.001 (0.001) (0.41)	0.001 (0.001) (0.40)	0.001 (0.001) (0.19)	0.001 (0.001) (0.27)	
MMBI × RBAN			−1.99 (2.63) (0.94)			
ECG/cap.	−1.16 (0.544) (0.03)	−0.20 (0.12) (0.12)				−0.01 (0.09) (0.87)
TERED	−0.14 (0.14) (0.34)	−0.10 (0.09) (0.26)	−0.02 (0.07) (0.78)	−0.01 (0.04) (0.08	−0.09 (0.05) (0.08)	0.44 (0.18) (0.02)
EMP_IND			−0.70 (0.22) (0.01)	−0.63 (0.24) (0.03)	−0.63 (0.20) (0.01)	0.21 (0.77) (0.77)
ED_SPEND			−0.39 (0.17) (0.04)	−0.46 (0.28) (0.10)	−0.21 (0.21) (0.33)	0.74 (1.74) (0.67)
SHADOW_EC	−0.45 (0.212) (0.03)	−0.21 (0.08) (0.03)			−0.16 (0.13) (0.27)	
SHADOW_EC × ECG/cap.	0.055 (0.02) (0.04)	0.007 (0.005) (0.21)				
Globalization						
INNOV		0.06 (0.03) (0.10)		0.05 (0.07) (0.53)	0.02 (0.06) (0.66)	−0.20 (0.23) (0.38)
HIGHTECHXP			0.04 (0.11) (0.68)			−0.76 (0.52) (0.10)
OPENESS	−0.02 (0.03) (0.33)	−0.004 (0.02) (0.86)	−0.003 (0.02) (0.88)			
Quality of institutions						
RULE_OF_LAW	−1.11 (3.33) (0.73)	0.61 (2.50) (0.81)	1.12 (2.09) (0.60)		0.89 (2.22) (0.69)	
CONT_CORR				−1.97 (1.19) (0.10)		
Sargan J-stat	41.44 (0.06)	66.91 (0.10)	65.26 (0.10)	61.11 (0.10)	61.80 (0.10)	0.56 (0.10)

Table 3. *Cont.*

	M1	M2	M3	M4	M5	M6
			Quality of institutions			
No. of instruments (groups)	10	10	10	10	10	10
Obs.no.	120	120	120	120	120	120

Note: Within the table, the coefficients are displayed together with standard errors and the probabilities within (). Standard errors are typically displayed in parentheses right below the coefficients to indicate they are related but distinct values.

5. Conclusions

The investigation into the effects of marketization on income inequality across CEE offers critical insights, synthesizing empirical evidence with dynamic panel data and GMM analysis. This refined understanding leads to several key conclusions:

The pronounced persistence of income inequality, as highlighted by the lagged income inequality variable's significance, underscores the chronic nature of disparities within CEE nations. This revelation underscores the imperative for enduring, strategic policy interventions designed to combat income inequality effectively.

Confirming the pivotal roles of economic growth and increased allocations for education, this study advocates for policies that bolster economic development while significantly investing in education. Such initiatives promise to foster equitable income distributions, enhance job quality, and broaden educational opportunities.

The nuanced impacts of minimum wage adjustments and the unequivocally positive influence of employment protection and active labor market initiatives illustrate the essential nature of thoughtful labor market policies. These findings advocate for measures that uplift low-income workers and promote inclusivity within the labor market.

The association between reduced income inequality with high technology exports and economic openness speaks to globalization's multifaceted role in fostering equitable income distribution. This highlights the opportunities globalization presents for inclusive growth, emphasizing strategic global integration.

This study illuminates the indispensable role of governance, with strong institutions marked by rule of law and anti-corruption efforts emerging as crucial for mitigating income disparities. Strengthening governance and institutional integrity is framed as a cornerstone strategy in addressing income inequality.

The shadow economy's complex influence on income inequality highlights the intricate balance needed in integrating the informal sector with the formal economy. Crafting strategies that harness the informal sector's potential while curbing its adverse effects is pivotal for equitable growth.

In sum, this study not only enriches the academic dialogue on income inequality within the context of CEE countries but also provides actionable insights for policymakers. By delineating the mechanisms through which marketization factors influence income distribution, it underscores the critical need for targeted, integrated policy interventions that span economic, educational, labor, and governance domains to cultivate a more inclusive, equitable economic landscape.

While this study contributes valuable insights into income inequality within CEE countries, certain limitations must be acknowledged. The exclusion of countries like Croatia due to insufficient data highlights the broader issue of data limitations in post-communist nations, emphasizing the challenges associated with comprehensive regional analyses. Additionally, despite the study's extensive coverage, it may not capture all relevant marketization factors and their nuanced interactions. Factors such as technological innovation, demographic shifts, and international trade dynamics warrant further exploration for a comprehensive understanding. Furthermore, the focus on CEE countries, while providing essential regional context, raises concerns about the generalizability of the findings to other global contexts. The unique historical, economic, and social trajectories of CEE nations underscore the need for caution when applying these results beyond the studied region.

Future research endeavors in the realm of income inequality within CEE countries could prioritize enhanced data collection efforts, particularly in nations currently facing data limitations. Additionally, future research could delve into evaluating the effectiveness of specific policy interventions within the CEE context, contributing to a more targeted and evidence-based approach to reducing income inequality in the region.

Social implications: The findings highlight the importance of inclusive growth that benefits a broader segment of the population. By addressing income inequality, countries can improve social cohesion, reduce poverty rates, and enhance the overall quality of life for their citizens. This approach aligns with the pursuit of the Sustainable Development Goals, particularly Goal 10, which focuses on reducing inequalities.

Economic implications: Addressing income inequality is not just a matter of social justice but also economic efficiency. High levels of inequality can hinder economic growth, create economic instability, and waste human capital. Policies that foster a more equitable income distribution can lead to a more sustainable and robust economic system.

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

Funding: This research study has been elaborated within the Data Science Research Lab for Business and Economics of the Bucharest University of Economic Studies.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data can be available upon request.

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

Abbreviations

Acronyms for the variable.

Acronym	The Name of the Variable
GINI COEF	Gini Coefficient
MMWBI	Minimum monthly salary, annual average
GDP/cap.	Gross domestic product per capita
ECG/cap.	Economic growth/cap
TERED	Employed population with tertiary education
EMP_IND	Employees in the industry
ED_SPEND	Education expenditure
INNOV	Innovation index
HIGHTECHXP	Share of high-tech exports
OPENESS	Openness of the economy
REG_QUAL	Regulatory quality
RULE_OF_LAW	Rule of law
CONT_CORR	Control of Corruption
SHADOW_EC	Informal economy
URB	Urbanization degree
EMP_SEV	Employees in the services sector

References

Acemoglu, Daron, and David Autor. 2011. Skills, tasks, and technologies: Implications for employment and earnings. *Handbook of Labor Economics* 4B: 1043–171.

Acemoglu, Daron, and James A. Robinson. 2012. *Why Nations Fail: The Origins of Power, Prosperity, and Poverty*. New York: Crown Business.

Atkinson, Anthony Barnes, and Salvatore Morelli. 2010. Inequality and labor market institutions. *Handbook of Income Distribution* 2: 1059–143.

Autor, David H., Lawrence F. Katz, and Melissa S. Kearney. 2008. Trends in U.S. Wage Inequality: Revising the Revisionists. *The Review of Economics and Statistics* 90: 300–23. [CrossRef]
Barro, Robert J. 2000. Inequality and Growth in a Panel of Countries. *Journal of Economic Growth* 5: 5–32. [CrossRef]
Berg, Andrew, and Jonathan D. Ostry. 2011. *Inequality and Unsustainable Growth: Two Sides of the Same Coin?* International Monetary Fund Staff Discussion Note No. 11/08. Washington, DC: International Monetary Fund.
Bergstrand, Jeffrey H., and Peter Egger. 2007. A Knowledge-and-Physical-Capital Model of International Trade Flows, Foreign Direct Investment, and Multinational Enterprises. *Journal of International Economics* 73: 278–308. [CrossRef]
Buehn, Andreas, and Friedrich Schneider. 2012. Shadow economies around the world: Novel insights, accepted knowledge, and new estimates. *International Tax and Public Finance* 19: 139–71. [CrossRef]
Card, David, and Alan B. Krueger. 1994. Minimum Wages and Employment: A Case Study of the Fast-Food Industry in New Jersey and Pennsylvania. *The American Economic Review* 84: 772–93.
Chetty, Raj, Nathaniel Hendren, Patrick Kline, Emmanuel Saez, and Nicholas Turner. 2014. Is the United States still a land of opportunity? Recent trends in intergenerational mobility. *American Economic Review* 104: 141–47. [CrossRef]
Firebaugh, Glenn, and Brian Goesling. 2004. Accounting for the Recent Decline in Global Income Inequality. *American Journal of Sociology* 110: 283–312. [CrossRef]
Forbes, Kristin J. 2000. A Reassessment of the Relationship between Inequality and Growth. *American Economic Review* 90: 869–87. [CrossRef]
Goldin, Claudia, and Lawrence F. Katz. 2007. *The Race between Education and Technology*. Cambridge: Harvard University Press.
Goos, Maarten, and Alan Manning. 2007. Lousy and Lovely Jobs: The Rising Polarization of Work in Britain. *The Review of Economics and Statistics* 89: 118–33. [CrossRef]
Gouvea, Raphael, and Wei Wang. 2019. Innovation and income inequality: Evidence from China's high-tech industry. *Structural Change and Economic Dynamics* 51: 133–44.
Kaufmann, Daniel, Aart Kraay, and Massimo Mastruzzi. 2010. The worldwide governance indicators: Methodology and analytical issues. *Hague Journal on the Rule of Law* 2: 220–46. [CrossRef]
Li, Xiaoyig, and Xiaming Liu. 2005. Foreign direct investment and economic growth: An increasingly endogenous relationship. *World Development* 33: 393–407. [CrossRef]
Lin, Frank, and John Li. 2011. Exports and income inequality: Further evidence from China. *China Economic Review* 22: 80–91.
Medina, Leandro, and Friedrich Schneider. 2018. *Shadow Economies around the World: What Did We Learn over the Last 20 Years?* Washington, DC: International Monetary Fund.
Milanovic, Branko. 2005a. Can we discern the effect of globalisation on income distribution? Evidence from household surveys. *The World Bank Economic Review* 19: 21–44. [CrossRef]
Milanovic, Branko. 2005b. *Worlds Apart: Measuring International and Global Inequality*. Princeton: Princeton University Press.
Murtin, Fabrice, and Romain Wacziarg. 2014. The democratic transition. *The Quarterly Journal of Economics* 129: 267–340.
Persson, Tabellini, and Guido E. Tabellini. 1994. Is Inequality Harmful for Growth? *American Economic Review* 84: 600–21.
Piketty, Thomas. 2014. *Capital in the Twenty-First Century*. Cambridge: Cambridge Belknap Press.
Rodrik, Dani. 1997. *Has Globalization Gone Too Far?* Washington, DC: Institute for International Economics.
Schneider, Friedrich, and Dominik H. Enste. 2000. Shadow Economies: Size, Causes, and Consequences. *Journal of Economic Literature* 38: 77–114. [CrossRef]
Torgler, Benno, and Friedrich Schneider. 2007. Shadow economy, voice and accountability, and corruption. In *Growth, Inequality, and Poverty: Prospects for Pro-Poor Economic Development*. Edited by Anthony F. Shorrocks and Rolph Van Der Hoeven. Oxford: Oxford University Press, pp. 153–71.

economies

MDPI

Article

From Policy to Impact: Advancing Economic Development and Tackling Social Inequities in Central and Eastern Europe

Adriana AnaMaria Davidescu [1,2,*], Tamara Maria Nae [3,4] and Margareta-Stela Florescu [5]

[1] Department of Statistics and Econometrics, The Bucharest University of Economic Studies, 010552 Bucharest, Romania
[2] Department of Education, Training and Labour Market, National Scientific Research Institute for Labour and Social Protection, 010643 Bucharest, Romania
[3] Department of Economics and Economic Policy, Bucharest University of Economic Studies, 010552 Bucharest, Romania; tamara.nae@economie.ase.ro
[4] Research and Analysis Department, Ministry of Finance, 050706 Bucharest, Romania
[5] Department of Administration and Public Management, Bucharest University of Economic Studies, 010374 Bucharest, Romania; margareta.florescu@ari.ase.ro
* Correspondence: adriana.alexandru@csie.ase.ro

Abstract: This study challenges the traditional reliance on GDP as the sole indicator of the success of the EU's cohesion policy, aligning with the evolving academic discourse that calls for a broader spectrum of metrics incorporating social factors. The research aims to assess the impact of cohesion on economic performance and social progress at the regional level in Central and Eastern European countries, using regression analysis on panel data. Inspired by the call to move beyond GDP-focused assessments, this research re-evaluates cohesion policy within an expanded framework that prioritizes economic and social dimensions. Specifically, it addresses the escalating concerns of income disparity and poverty in Central and Eastern European nations. Utilizing panel data regression models, this study scrutinizes data from 2007 to 2018, covering two recent programming periods, to offer a comprehensive, multifaceted analysis of the impact of cohesion policy. It underscores the policy's dual role in spurring economic growth and fostering social progress, particularly in mitigating income inequality and reducing poverty. The findings reveal that cohesion policies positively affect both economic performance and social progress, with notable impacts on narrowing the income gap and alleviating poverty in these regions. However, the economic benefits for poverty reduction materialize over a prolonged period, reflecting the gradual nature of policy impact and the time needed for investments to materialize. The study emphasizes the need for a long-term strategic vision in implementing cohesion policies. This includes enhanced data collection, a deeper focus on the social ramifications of policies, streamlined policy processes, capacity building, institutional strengthening, and prioritizing equitable opportunities to bridge income gaps effectively. This comprehensive approach aims to maximize the dual benefits of cohesion policies, promoting balanced economic and social progress across Central and Eastern Europe.

Keywords: cohesion policy; economic development; social progress; regional analysis; beyond the GDP

Citation: Davidescu, Adriana AnaMaria, Tamara Maria Nae, and Margareta-Stela Florescu. 2024. From Policy to Impact: Advancing Economic Development and Tackling Social Inequities in Central and Eastern Europe. *Economies* 12: 28. https://doi.org/10.3390/economies12020028

Academic Editor: Fabio Clementi

Received: 5 December 2023
Revised: 17 January 2024
Accepted: 19 January 2024
Published: 24 January 2024

1. Introduction

The final objective of the European Social Model is to simultaneously ensure economic growth and social cohesion. One of the central goals of the EU (European Union), stipulated in the Maastricht Treaty, is promoting social and economic progress by strengthening economic and social cohesion. It is widely recognized that the promotion of cohesion is one of the most prominent and important of the EU's many political responsibilities; this importance is founded on several aspects, such as: (i) the fact that it has acquired increased importance over time with regard to the budgetary expenses of the European Union;

(ii) cohesion policy is quite visible in its broad remit, covering a very wide range of EU policy activities, including infrastructure, telecommunications, research and development, competitiveness, vocational training, employment and social inclusion, as well as objectives to promote environmental sustainability and digitization—objectives that are in line with the perspectives of the green and digital transition; (iii) the contribution to the consolidation of the process of historical enlargement and the EU deepening; (iv) cohesion policy involves a large number of political actors at European level, also managing to include governmental and non-governmental decision-makers at regional and local levels.

With the emergence of the Europe 2020 Strategy, the objectives of the cohesion policy have acquired a multidimensional characteristic, paying increasing attention to social objectives. Thus, the analysis to be developed, as part of quantitative research, aims to expand the assessment of the cohesion policy on economic development by adding the social dimension to identify the effects in terms of achieving the two objectives that are complementary in the European model: economic growth and social cohesion.

This study explores the hypothesis that cohesion policy positively affects economic growth and social outcomes. The research probes into the impact of cohesion policy on economic performance and social progress by analyzing regional-level data from Central and Eastern European (CEE) countries during the 2007–2013 and 2014–2020 programming periods using panel data regression models. This choice stems from the EU's financial allocation strategy based on regional GDP (gross domestic product) and the increasing recognition of the social dimension in both EU and global strategies, like the UN's Sustainable Development Goals.

Central and Eastern European countries exhibit diverse economic performances and social landscapes. Before the COVID-19 pandemic, many countries like Poland, Hungary, Czech Republic, Slovakia, and Romania experienced GDP growth rates between 3% and 5% annually, reflecting relative economic robustness compared to Western Europe. However, varying levels of unemployment persist, with rates around 6.1% in Poland, 4.1% in Hungary, 3.6% in the Czech Republic, 7.7% in Slovakia, and 4.3% in Romania as of 2021. Income disparities remain a challenge, particularly between rural and urban areas, accentuating economic inequalities. While some countries have improved healthcare systems and social welfare, demographic shifts, including aging populations and the outward migration of skilled workers, pose ongoing concerns. EU funding has aided infrastructure development and education, and digital transformation efforts are underway to bolster innovation and competitiveness. The COVID-19 pandemic has brought varied economic impacts to CEE nations. Some experienced significant contractions due to lockdowns and disruptions in global supply chains, while others demonstrated resilience. These countries face ongoing challenges in managing income disparities, unemployment, and demographic shifts while striving to modernize infrastructure, improve healthcare, and enhance digital capabilities.

A defining feature of the CEE nations that justifies grouping them for analysis, anticipating relatively consistent outcomes, is their shared history as former communist states. These nations underwent challenging transitional phases, leaving them with various economic, social, and institutional vulnerabilities. Consequently, their administrative capabilities in securing external financial support and their methods of utilizing these funds have not achieved the anticipated progress levels. The ERDF (European Regional Development Fund) stands out as one of the most pivotal tools, championing goals like economic expansion and employment generation and fostering close territorial collaboration. Additionally, the ERDF lends support to regions facing inherent demographic or geographical challenges. This includes regions like thinly populated areas or those dominated by mountainous terrains where homes are widely dispersed from the main community. For the 2014–2020 programming duration, out of the total budget earmarked for the cohesion policy (around EUR 350 billion), the ERDF was allocated EUR 199 billion, earmarked for specific thematic goals.

This study aims to extend the assessment of cohesion policy's impact on economic development by adding a social dimension, thus exploring the dual objectives of economic

growth and social cohesion, which are complementary in the European model. Thus, the study aims to respond to several relevant research questions:

How does the EU's cohesion policy impact economic growth and social outcomes in Central and Eastern European (CEE) countries? How effective are EU cohesion policies in reducing social inequities, specifically income disparity and poverty, in CEE regions?

Is there a measurable correlation between the implementation of EU cohesion policies and improvements in both economic performance and social progress at the regional level in CEE countries? How does the time-lagged nature of cohesion policy's impact influence its effectiveness in achieving economic and social objectives?

However, it should be noted that (which is characterized by an alternation of starting and stopping the reforming processes) (Dinu et al. 2005), which is why one of the methodological difficulties identified at the level of this analysis is that the effects of cohesion policy on economic performance and social progress cannot be observed immediately but with delays. Therefore, the effects of cohesion policy take time to become apparent as investments are spread over several years, and some of their results are then reinvested, which delays the expected effects. At the same time, some regions are more dynamic in economic activity, and others have a narrower concentration of activity, which determines different rhythms in terms of the emergence of results. Thus, the structure and implementation rules of the cohesion policy made it vulnerable to criticism (Barca 2009). Hence, there is a need to identify the efficiency or inefficiency of the cohesion policy as a contribution to the discussions on its reformation. Thus, in the framework of the analysis, models with a dynamic structure were developed to capture the delays in the appearance of the results. In enhancing the discourse, this study pushes boundaries beyond traditional GDP-focused evaluations, advocating for more holistic indicators that encapsulate social progress. This endeavor augments the academic dialogue and informs policy decisions for upcoming programming periods. A review of the existing literature suggests a pressing need to explore cohesion policy's dual impact on economic development and social progress. Moving beyond the traditional metric of GDP growth, there is a compelling case for incorporating broader indicators that offer a comprehensive insight into social dimensions. While GDP growth remains a vital benchmark, it is not the sole aim of cohesion policy. Evaluating its effects on diverse socio-economic markers deepens academic discourse and bolsters the underpinnings of policy decisions. Such a holistic approach could enhance the efficiency of cohesion fund allocations in upcoming programming periods. Cohesion funds are EU subsidies for the development of the social and industrial infrastructure of certain member nations.

Incorporating social indicators, such as the SPI (Social Progress Index), into the evaluation of cohesion policy is crucial as it provides a more comprehensive understanding of policy impacts; traditional economic metrics like GDP are limited in scope and fail to fully capture the nuances of societal advancement, especially in the realms of income inequality and poverty reduction.

The Opportunities pillar of the SPI, focusing on aspects like access to education, information, and advanced healthcare, serves as a vital complement to economic data, offering a deeper insight into how cohesion policies foster equitable opportunities and directly contribute to mitigating social disparities.

By integrating robust social indicators alongside economic metrics, we can gain a more holistic view of cohesion policy outcomes, ensuring that assessments are not just about economic efficiency but also about their effectiveness in creating inclusive, equitable societies where opportunities for advancement and poverty reduction are realistically appraised and addressed.

This study explores the hypothesis that cohesion policy positively affects economic growth and social outcomes. It examines the impact of cohesion policy on economic performance and social progress at the regional level in CEE countries during the 2007–2013 and 2014–2020 programming periods using panel data regression models. This approach is rooted in the EU's financial allocation strategy based on regional GDP and acknowledges

the increasing recognition of the social dimension in global strategies, such as the UN's Sustainable Development Goals.

CEE countries present diverse economic performances and social landscapes, with various challenges and achievements in different sectors. The COVID-19 pandemic has further complicated these scenarios, having varied economic impacts on the nations. The shared history of these countries as former communist states provides a common thread for analysis, anticipating relatively consistent outcomes across the region. From this point of view, the paper makes several contributions to the literature. Firstly, a holistic evaluation of cohesion policy: While much of the existing literature focuses on the economic impacts of the cohesion policy, primarily gauging success through GDP growth, this study provides a more comprehensive assessment; by incorporating the social dimension, the research presents a multidimensional perspective on the policy's outcomes. Secondly, an in-depth regional analysis of CEE countries: The study delves deep into the specific context of these countries. Given that the regions of CEE countries have traditionally been underrepresented in similar research, this work fills a significant gap. Thirdly, dynamic models: recognizing the time-lagged impacts of cohesion policies, the study employs dynamic models to account for the delays in outcomes. This approach ensures a more accurate depiction of the policy's real-time and long-term effects. Fourthly, the study acknowledges and validates other significant indicators reflecting social progress by venturing beyond the conventional GDP metric. This broader approach offers a more rounded picture of the policy's impacts and underscores the need for varied evaluation metrics in the future. Fifthly, implications for policy reformation: highlighting the cohesion policy's efficiency or inefficiency contributes valuable insights that could shape discussions on policy reformation. Sixthly, bridging literature gaps: the research takes a bold step in analyzing and challenging the existing literature. It addresses the dearth of studies on the social impact of cohesion funds, further enriching academic discussions and offering fresh avenues for future research.

This paper contributes to the literature by providing a holistic evaluation of cohesion policy, delving into the specific context of CEE countries, employing dynamic models to capture time-lagged impacts, and advocating for a broader range of evaluation metrics. It addresses the need for policy reformation and aims to bridge literature gaps by analyzing the social impact of cohesion funds.

The paper unfolds as follows: Initially, we outline our core objectives and the rationale behind the impact analysis, grounding our research within the European landscape. Next, we provide a literature review, spotlighting the scant research on the effectiveness of cohesion policy, particularly in the context of CEE regions. In the subsequent section, we delve into our research methodology, detailing the data transformation processes to ensure accurate outcomes. Our findings highlight the correlation between cohesion funds, economic growth, and social progress indicators. We then evaluate the recent evolution of cohesion policy tools in CEE over the past two programming periods. The paper concludes with key takeaways and actionable recommendations for enhancing the efficacy of cohesion fund allocation.

2. Literature Review

Over time, cohesion policy has proven effective in helping many states reduce their regional gaps and gaps with other Member States. It substantially improved the economies of the states in the PIGS group (Portugal, Ireland, Greece and Spain), having a GDP per inhabitant value below 90% of the European Community average, thus being eligible for the granting of cohesion funds and subsequently enjoying a substantial catching-up effect. Through its actions, cohesion policy has significantly contributed to accelerating growth and prosperity in the EU, reducing certain economic, social and territorial disparities.

In the last few decades, Fiaschi et al. (2018) and similar studies have provided valuable insights into the economic dimensions of cohesion policy. They predominantly center on regional GDP, neglecting a comprehensive analysis of crucial social outcomes such as employment rates, poverty levels, and social inclusion measures, thereby leaving a

notable gap in understanding the full spectrum of policy effects. The prevalent focus on economic indicators in the existing literature overlooks the multifaceted nature of social progress, failing to adequately evaluate how cohesion policies influence more nuanced social dimensions, including the effectiveness of these policies in directly reducing poverty, enhancing job opportunities, and fostering inclusive social environments. This oversight in the literature highlights a critical need for expanded research that goes beyond economic metrics, offering a more holistic view of cohesion policy impacts by incorporating assessments of direct social benefits, thereby providing a clearer picture of how these policies contribute to or fall short in addressing key societal challenges like income inequality and social exclusion.

In addition to the merits associated with cohesion policy, it has been subjected to a series of criticisms over time, such as: (i) failure to meet the objectives established by the EU Treaties, (ii) insufficient emphasis on economic growth (Sapir 2003), and (iii) turned into a captivating policy without a clear mission, with complex and overly bureaucratic administration (Manzella and Mendez 2009). Later, criticism narrowed down to the results, with studies identifying a positive impact on economic growth (Crescenzi and Giua 2016; Rodriguez-Pose and Novak 2013; Tomova et al. 2013). In contrast, other studies identified a negative impact (Dall'Erba and Le Gallo 2008) or no impact (Boldrin and Canova 2001), and others suggested that there is no consensus regarding the effectiveness of cohesion policy (Ederveen et al. 2003; Darvas and Wolff 2018). There are divergent conclusions about the effectiveness of cohesion policy because the methods and hypothetical approaches used in determining the results are extremely different.

Analyses of the impact of cohesion policy on European regional performance mainly focus on the economic dimension, measured by GDP per capita and occasionally by the employment rate (Becker et al. 2010; Rodríguez Pose and Novak 2013; Giua 2017; Fiaschi et al. 2018; Crescenzi and Giua 2020) or the level of education and health (Calegari et al. 2021). Although there is a rich literature base on the effects of cohesion, especially on economic performance, it is inconclusive. Contrasting results derive from the choice of spatial and temporal considerations, variables used, and impact estimation methodologies.

Farrell (2004) identifies a positive impact of structural funds on regional economic growth in Ireland and Spain. The results are supported by Lolos (2009), who analyzes the case of Greece. Some studies identify a positive impact on economic growth, with more pronounced effects for developed regions (Crescenzi and Giua 2016; Calegari 2020). The possible explanations for these results are, on the one hand, that the regions behind have less negotiation skills to attract more funds (Charron 2016; Fratesi and Wishlade 2017), and on the other hand, that these regions present a reduced capacity to absorb the allocated funds, which leads to the paradoxical situation where the regions entitled to receive considerable amounts from the structural funds cannot use them (Becker et al. 2013; Surubaru 2017; Cerqua and Pellegrini 2018).

Also, some studies have not identified any effect of cohesion policy on economic growth performance (Ederveen et al. 2006), which does not identify "conclusive agreements on the impact of EU cohesion policy in the existing literature" Medeiros (2014). Other studies suggest that significant transport and infrastructure investments supported by cohesion policy do not affect economic growth (Crescenzi and Rodríguez-Pose 2012). Some conclude that there is no significant impact of the absorption rate of EU funds on growth in the EU countries in the short term. However, even a negative impact can be identified (Albulescu and Goyeau 2013). Other studies identify a negative impact, such as Dall'Erba and Le Gallo (2008), who evaluated the impact of structural funds on the convergence process between 145 European regions from 1989 to 1999. They identified the convergence process, but the funds did not impact it. Although studies on the impact of cohesion funds have focused on the effects exerted on economic growth, with a limited contribution to the literature regarding the influence on other indicators that reflect the social dimension, recently, an increasingly significant wave of research explores the effect of cohesion policy on other types of non-economic outcomes as well (Ferrara et al. 2022; Albanese et al. 2021).

For example, Calegari et al. (2021) goes beyond GDP and analyzes the impact of cohesion policy on GDP per capita and societal well-being through a modified version of the adapted human development index. The results of the study indicate that cohesion policy has significantly increased overall well-being in low-performing regions that have used cohesion funds, with the results being particularly visible in improving the level of education. Attempts to overcome the GDP barrier were also approached by Crescenzi and Giua (2018), who found that the cohesion policy exerted a positive and significant impact at the EU level both on regional economic growth and on employment, suggesting that the positive effect on regional employment survived the Great Depression and supported less developed regions during the recovery.

Maucorps et al. (2020) analyzed the effects of EU cohesion policy on the economic growth of 276 European NUTS-2 regions between 2008 and 2016, using a structural equation model consisting of a measurement component (with two latent variables) and a structural component. The study's results support the existence and purpose of cohesion policy, where EU funding is essential for the economic development of European regions without other abundant sources of funding, focusing on mitigating structural deficiencies that prevent the effective use of convergence investments.

It is well known in the literature (Acemoglu and Robinson 2012) that economic growth, often stimulated by effective cohesion policies, lays the foundation for social progress, as increased regional prosperity can lead to enhanced public services, better infrastructure, and improved living standards, all of which are crucial for comprehensive social development. While economic advancements under cohesion policy have the potential to yield significant social benefits, such as reduced poverty rates and greater income equality, there remains a conspicuous gap in the current research that explicitly links these economic gains to tangible improvements in these specific social conditions (Stiglitz et al. 2010).

According to Piketty (2014), the assumption that economic growth automatically translates into social progress is overly simplistic; without deliberate measures and targeted policies, the benefits of increased regional GDP may not effectively trickle down to address core issues of poverty and income disparity. To truly understand the impact of cohesion policy, research must extend beyond economic indicators and rigorously examine how these economic improvements correlate with, and possibly contribute to, key social outcomes like enhanced employment opportunities, poverty alleviation, and narrowed income gaps (Sen 1999).

This lack of explicit linkage in the current literature between economic growth and its potential social benefits underscores the need for a more integrated approach to policy evaluation—one that considers how economic advancements under cohesion policies are practically reflected in the daily lives of individuals, particularly those in economically disadvantaged segments (Sachs 2015).

While the literature has established the economic effectiveness of cohesion policy, particularly for the PIGS countries, there is a discernible need to delve into the social impacts of such policies. The current research corpus, which predominantly centers on GDP growth, does not sufficiently address the broader social dimensions such as employment, poverty alleviation, and social inclusion, leaving a gap our research aims to fill. We propose a dual-focused analysis considering economic and social indicators to provide a more nuanced understanding of cohesion policy outcomes. The contrasting conclusions drawn from existing studies on the economic impact of cohesion policy, ranging from positive to non-significant effects, suggest a methodological divergence that our study seeks to reconcile. By employing a robust, integrative methodology, we aim to offer clarity to the debate and contribute a comprehensive perspective on the effectiveness of cohesion policy in both economic and social terms.

Furthermore, while some research points to the benefits of cohesion policy on regional development and growth, the social implications of such policies are not as well documented. Our study extends the scope of analysis beyond GDP, incorporating broader indicators of societal well-being, such as the modified HDI (Human Development Index),

as examined by Calegari et al. (2021), and employment, as highlighted by Crescenzi and Giua (2018).

We acknowledge the literature's call for an integrated evaluation approach that tracks economic progress under cohesion policy and critically examines how such advancements translate into social benefits. By doing so, we aim to bridge the gap identified by seminal thinkers like Stiglitz et al. (2010) and Piketty (2014) and provide a more holistic assessment of cohesion policy's impact on improving living standards and reducing social disparities.

In essence, our research addresses the lacuna in the existing literature by exploring the direct linkage between economic growth facilitated by cohesion policy and its tangible social benefits, offering new insights into how increased regional prosperity under such policies can lead to substantial social improvements, particularly for those in less advantaged economic brackets.

To contextualize our study within the broader academic dialogue and to illustrate the contribution this research makes to existing scholarship, Table 1 below presents a curated overview of major previous studies related to the impact of cohesion policy on economic growth and social outcomes.

Table 1. An overview of the most relevant studies in the literature.

Authors	Title	Journal	Year
Fratesi, Ugo, and Fiona G. Wishlade	The Impact of European Cohesion Policy in Different Contexts	Regional Studies	2017
Gagliardi, Luisa, and Marco Percoco	The Impact of European Cohesion Policy in Urban and Rural Regions	Regional Studies	2017
Crescenzi, Riccardo, and Mara Giua	One or Many Cohesion Policies of the European Union? On the Differential Economic Impacts of Cohesion Policy across Member States	Regional Studies	2020
Bradley, John	Evaluating the Impact of European Union Cohesion Policy in Less-Developed Countries and Regions	Regional Studies	2006
Pîrvu, Ramona, et al.	The Impact of the Implementation of Cohesion Policy on the Sustainable Development of EU Countries	Sustainability	2019

The existing literature on cohesion policy in the EU has predominantly concentrated on economic indicators, such as regional GDP, neglecting crucial social dimensions like employment rates, poverty levels, and social inclusion. To address this gap, our proposed research takes a dual-focused approach, simultaneously considering economic and social indicators. In doing so, we aim to offer a more comprehensive and holistic understanding of the outcomes of cohesion policy. Additionally, there is a recognized methodological divergence in existing studies regarding the economic impact of cohesion policy. Our research contributes by proposing the use of a robust, integrative methodology to reconcile these differences, providing a comprehensive perspective on the effectiveness of cohesion policy in both economic and social terms. This approach aims to bridge the gap in the literature and enhance the understanding of the multifaceted impacts of cohesion policies in the European Union.

3. Data and Methodology

This research aims to assess the impact of cohesion on economic performance and social progress at the regional level in CEE countries using regression analysis on panel data. The study focuses on 54 regions of 10 CEE countries (Bulgaria, Czech, Estonia, Latvia, Lithuania, Hungary, Poland, Romania, Slovakia and Slovenia), for the period 2007–2018—this period was chosen according to the last two programming periods of the Multiannual

Financial Framework (2007–2013–2014–2020). Still, data for cohesion funds at the regional level were available until 2018. Cohesion funds at the regional level were accumulated for the years in which the payments related to the two programming periods overlapped (according to the rules for the allocation of European funds and after the completion of the programming period. The variables related to the allocations of cohesion funds—ERDF, CF (Cohesion Fund), and ESF (European Social Fund)—were expressed according to the number of inhabitants for each region and deflated using the GDP deflator with the fixed base 2015 = 100. A detailed description of the variables used in the model and the source of the data is presented in Appendix A, Table A1.

To evaluate the impact of cohesion instruments on economic development and social progress, several specifications have been evaluated using two main result indicators: regional GDP and the social progress index and its pillars. To do that, a logarithmic regression model, represented in a log-log form, has been used, allowing for elasticity interpretation, meaning a percentage change in the independent variable results in a percentage change in the dependent variable, acknowledging four different specifications, one for each type of fund:

$$\log(real\ GDP/cap)_{it} = \beta_0 + \beta_1 \cdot \log\left(invest.\frac{ESIF}{cap}.\right)_{it-1} + \beta_2 \cdot \log(\text{initial}.\frac{\text{GDP}}{\text{cap}}.)_{it} + \beta_3 \cdot EQI_{it} + \varepsilon_{it} \quad (1)$$

where i = 1 … 54 represents the 54 regions of the 10 CEE countries; t = 1, 2, … T represents time (2007–2018), and β_i are the model parameters; real GDP/cap. was deflated using the GDP deflator with the base 2015 = 100; the ESIF investments/cap. from the previous period are expressed in comparable prices using the GDP deflator with the base 2015 = 100 and are calculated as the sum of investments from the ERDF, CF and ESF funds; and EQI is the governance quality index at the European level. The value of real GDP was taken from the National Accounts database of Eurostat, and the EQI was taken for three years from the surveys in 2010, 2013, and 2017 from the studies of Charron et al. (2014) and Nistotskaya et al. (2015). The initial value of real GDP per capita was determined by capturing the convergence effect. According to Maucorps et al. (2020), this variable refers to the phenomenon of spatial propagation, when a change in a region's investment has effects not only on the region itself but also on the economic development of other regions. Developments from other regions (caused by the initial change of the first investment in the region) are fed back into the original region. Therefore, these spatial feedback loops should be integrated into assessing the effects of cohesion policy on regional economic development to capture both the investment's direct and indirect (feedback) effects. In the present analysis, we use three specifications of the model—one model for each type of investment (ERDF, CF and ESF)—and another specification considering their aggregated value to decompose the impact into components.

It is important to mention that in all the models developed, the instruments of the cohesion policy—the funds—were introduced with a lag of a period, assuming that the effects of the cohesion policy manifest themselves with a delay.

Within this particular type of (log-log) model, the partial regression coefficients play the role of elasticities of the independent variable Xi concerning Y, that is, the percentage change in Y as a result of the percentage change in Xi, ceteris paribus.

To capture the impact of ESIF investments on social progress, we used a lin-log model, which allowed four different specifications, one for each type of fund:

$$SPI_{it} = \beta_0 + \beta_1 \cdot \log(inves.FSIE/cap.)_{it-1} + \beta_2 \cdot \log(initial\ real\ GDP/cap.)_{it} + \beta_3 \cdot EQI_{it} + \varepsilon_{it} \quad (2)$$

where i = 1 … 54 represents the 54 regions of the 10 CEE countries, t = 1, 2, …, T represents time (the period 2007–2018), and β_i are the model parameters; SPI is the social progress index, and (Basic Human Needs, Foundations of Well-being and Opportunities) are the components of SPI; the ESIF investments/cap. from the previous period are expressed in comparable prices using the GDP deflator with the base 2015 = 100 and are

calculated as the sum of investments from the ERDF, CF and ESF funds; and the EQI is the governance quality index at the European level.

The SPI is the most comprehensive measure of social progress. Another attempt to assess social progress "going beyond GDP" was the HDI. Still, the consistency and coverage of the SPI are superior (it contains 54 indicators grouped into 3 dimensions). Moreover, economic indicators are excluded from the SPI to allow for comparison with indicators such as GDP per capita. Furthermore, for this type of model, several specifications (three in number) were estimated that depend, however, on the change of the dependent variable (the SPI and its components: Basic Human Needs, Basic Well-being and Opportunities).

The main methodological Constraints have been:

- For the SPI, we utilized the 2011 figures for the 2007–2010 timeframe due to the data for this metric being only accessible for 2011–2020. Additionally, national SPI values were applied to regional levels, given that regional-level SPI data were only available for 2016 and 2020.
- The EQI data for the analysis span consist of three distinct sets. In instances where specific years were not directly covered, we adopted the value from the closest available year.

Estimating panel data models involves a systematic approach. After model specification, the Hausman test is a popular diagnostic tool

$$H_0 : \sigma_\alpha^2 = 0$$

(i.e., there is no individual-specific error variance)

Against the alternative

$$H_1 : \sigma_\alpha^2 > 0$$

The test statistic is given by

$$LM = T \times R^2 \tag{3}$$

from the auxiliary regression of the squared residuals from a pooled OLS regression on the individual means of the independent variables. If the LM statistic is significant, then heteroscedasticity (and correlation) is present across the individual-specific errors.

The Modified Wald Test for Groupwise Heteroskedasticity checks for differences in the error variance across groups (or cross-sectional units). Given the model of Equation (4):

$$yit = X_{it}\beta + \alpha_i + u_{it}$$

The null hypothesis is:

$$H_0 : = \sigma_1^2 = \sigma_2^2 = \ldots = \sigma_N^2$$

(i.e., the error variance is constant across groups)

Against the alternative:

H_1. *At least one σ_i^2 is different.*

The test statistic follows a chi-square distribution. If it is significant, it suggests heteroskedasticity across units.

Serial correlation or autocorrelation is the correlation of a variable with its past values. This is especially concerning in panel data because it can cause coefficient estimates to be inefficient. Serial correlation can also lead to underestimated standard errors, increasing the risk of Type I errors. Two tests are applied to test serial correlation: The Wooldridge and the Arellano–Bond tests. The first test is used for detecting first-order autocorrelation in panel data settings. Given the following model

$$y_{it} = X_{it}\beta + \alpha_I + u_{it} \tag{4}$$

the focus is on the residuals $\hat{u}_{it}$ from the fixed effects regression. The test regression is:

$$\hat{u}_{it} = \rho\hat{u}_{i,t-1} + \nu_{it}$$

The null hypothesis is:

$$H_0 : \rho = 0$$

(i.e., there is no first-order autocorrelation)

Against the alternative:

$$H_1 : \rho \neq 0$$

The test statistic is t-statistic for ρ in the test regression. If the test statistic significantly differs from zero, there is evidence of first-order auto-correlation.

This Arellano–Bond test is usually applied in the context of dynamic panel data models estimated using the Arellano–Bond GMM estimator. The test checks for autocorrelation in the first-differenced errors, which, if present, can indicate that the instruments used in the GMM estimation are not valid.

Cross-sectional dependence (CSD) in panel data refers to the likelihood that the errors associated with one entity (a country in a dataset of multiple countries) might be correlated with the errors of another entity. This phenomenon is especially prevalent in panels where entities might be influenced by unobserved common factors or where spillover effects between entities exist. Cross-sectional dependence can lead to biased and inconsistent coefficient estimates if not accounted for.

1. Pesaran's CD Test:

One common test for cross-sectional dependence is the Pesaran CD test. The test statistic is defined as:

$$CD = \sqrt{\frac{2}{T(N-1)}}\Sigma_{i=1}^{N-1}\Sigma_{j=i+1}^{N}\rho_{ij} \quad (5)$$

where:

- T is the number of periods.
- N is the number of cross-sectional units.
- ρ_{ij} is the sample correlation of the residuals from two cross-sectional units, i and j.

CD is asymptotically normally distributed under the null hypothesis of no cross-sectional dependence.

2. Breusch–Pagan LM Test:

Another popular test is the Breusch–Pagan LM Test for cross-sectional dependence. The test statistic is:

$$LM = T\left(\hat{p}^2 - \hat{q}^2\right) \quad (6)$$

where:

- T is the number of periods.
- $\hat{p}$ is the estimated average of the squared pairwise sample correlation of the residuals.
- $\hat{q}$ is the average of the squared sample correlations under the null of no cross-sectional dependence.

The R^2 value is a statistic that provides information about the goodness-of-fit of a model. In panel data regressions, we encounter three types of R^2: within, between, and overall. Each of these offers different insights into the model's fit at different data dimensions.

Within R^2 indicates how well the model explains the variation of the dependent variable around its mean within each cross-sectional unit (e.g., each individual, each country):

$$R^2_{within} = 1 - \frac{SSR_{within}}{SST_{within}} \quad (7)$$

SSR_{within} is the sum of squared residuals from the within transformation, and SST_{within} is the total sum of squares from the within transformation.

Between R^2 shows how well the model explains the variations of the dependent variable around its mean between the cross-sectional units, using the time-averaged values of the variables:

$$R^2_{between} = 1 - \frac{SSR_{between}}{SST_{between}} \quad (8)$$

where $SSR_{between}$ is the sum of squared residuals using time-averaged values, and $SST_{between}$ is the total sum of squares using time-averaged values.

Overall R^2 combines both within and between variations to measure how well the model explains the variation of the dependent variable around its grand mean:

$$\mathrm{R^2overall} = 1 - \frac{SSR_{overall}}{SST_{overall}} \quad (9)$$

where $SSR_{overall}$ is the total sum of squared residuals from the pooled regression (without any transformation), and $SST_{overall}$ is the total sum of squares without any transformation.

Each type of R^2 provides a unique perspective on the model's fit:

- Within R^2, it gives insights into how well our model explains variations over time for each entity.
- Between R^2 tells us how well our model explains differences between entities.
- Overall, R^2 provides a combined measure, reflecting the model's ability to explain both time series and cross-sectional variations in the panel data.

4. Empirical Results

4.1. Analysis of the Main Developments of Cohesion Policy Instruments at the Level of CEE Countries from the Perspective of the Last Two Programming Periods

At the level of the CEE states, Poland has benefited from the largest allocation to date, approximately EUR 49 billion, followed by the Czech Republic (EUR 17 billion) and Romania (EUR 14 billion). It should be noted that Poland is the CEE country with the highest performance in attracting European financial resources for each instrument. This performance of Poland is due to the implementation a transition model based on shock therapy, which proved to be a success, and to other reforms, such as the administrative one, which involved a territorial division allowing regional self-governance.

The cohesion fund is one of the main pillars of the cohesion policy. For 2014–2020, it was allocated according to the GNI (Gross National Income) per capita criterion to the Member States with a GNI per capita lower than 90% of the EU average. All the states from the CEE group were in this category, plus Cyprus, Malta and Greece. The financial resources allocated through the CF were approximately EUR 63.4 billion, distributed to the member states to make investments in the environment, sustainable development of transport infrastructure, and technical assistance. Poland, the Czech Republic and Romania are among the main beneficiaries.

The ESF finances both objectives aimed at economic cohesion, but it mainly deals with the social dimension of cohesion. The more than EUR 80 billion from the ESF has been allocated to all EU regions to improve aspects related to human capital, such as education and professional training; reduce the risk of poverty; or protect vulnerable people, so the ESF aims to invest in increasing the level of employment on the labor market and improving the skills of employees.

The highest level of ESF allocation is recorded by Poland (EUR over 15 billion), followed by Hungary and Romania with EUR 6 and 5.67 billion, respectively. Given these considerable sums allocated to the CEE countries, the question of its use and benefits to the respective countries regarding economic development and social progress is inevitable.

The first attempt to measure well-being beyond GDP was the HDI. Pakistani economist Mahbub ul-Haq developed the HDI based on Amartya Sen's considerations of human capabilities. According to Klugman (2011), the HDI was constructed to measure a country's

level of development to "shift the focus of economic development from national income accounting to people-centred policies"[1]. It is a summary measure of average achievement in key dimensions of human development: a long and healthy life, knowledge, and a decent standard of living. The HDI is the geometric mean of the normalized indices for each of the three dimensions[2].

Regarding its evolution in the period 2014–2020, it can be observed that all CEE countries have registered progress, with the biggest jump being made by Poland. Poland also benefited from the largest amount of funding from cohesion funds during this period. We believe that the explanation for Poland's progress can affect the effective use of cohesion funds.

According to Figure 1, for Romania, the most recent HDI value is 0.828, a value that places Romania in the category of very high human development, being in 49th place out of the 189 countries for which the HDI is calculated.

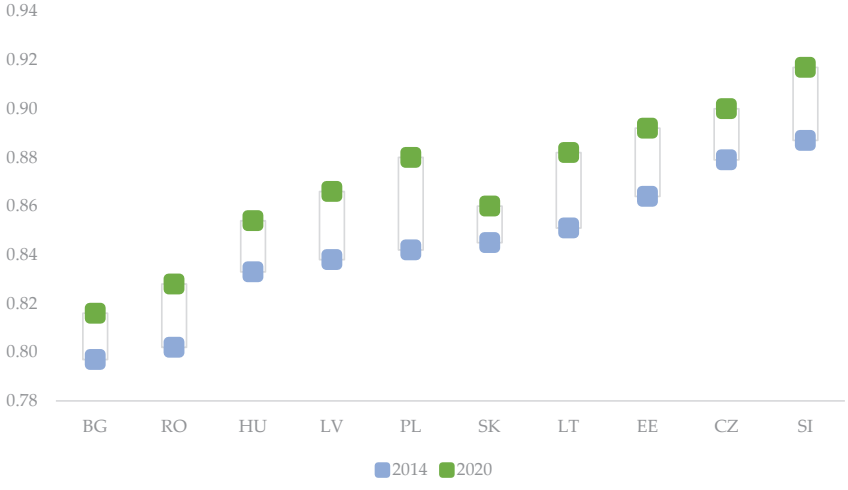


Figure 1. Evolution of the Human Development Indicator (2014–2020).

The SPI was developed based on the theories of Amartya Sen, Douglass North and Joseph Stiglitz[3] and measures citizens' quality of life and social well-being in 163 countries based on the analysis of 3 main dimensions consisting of 53 indicators. The methodology for its development consists of assigning a score for the items in the Basic Human Needs category (nutrition and basic health care, water and sanitation, shelter and personal safety), the Basic Well-being category (access to basic knowledge, access to information and communications, health and well-being, environmental quality), and the Opportunity category (personal rights, personal freedom and choice, inclusion, access to advanced education).

Going beyond GDP to reflect better societal development, the SPI 2020 shows that social disparities still vary widely between EU countries. Northern countries perform quite well, while Southeastern countries lag. Most EU regions perform well on Basic Human Needs such as nutrition and basic health care, water and sanitation, shelter and personal security.

Results vary regarding intermediate indicators of social development, such as access to basic knowledge, access to information and communication, health and well-being, and environmental quality. The biggest differences are related to indicators of opportunity, i.e., personal rights, personal freedom and choice, tolerance and inclusion, access to advanced education, and lifelong learning.

In Figure 2, we analyzed the SPI at the national level, taking into account the two best-performing countries (Sweden—91.32 and Finland—92.11) and the two least-performing countries (Romania—77.67; Bulgaria—78.38) from EUR EU in terms of their SPI score for

2020. The Social Progress Index shows that countries with higher social progress have been more resilient to the wider social impact of the COVID-19 crisis. The pandemic has brought trauma and devastation to communities around the world beyond the number of illnesses and deaths. It has affected mental health, food security, education and more. According to data from the 2021 Social Progress Index Report, countries with higher social progress are also those that have been most resilient to the non-health stress caused by the pandemic. Higher levels of social progress across the board, not just investment in a particular area, can be an important factor in resisting the multidimensional effects of future shocks[4].

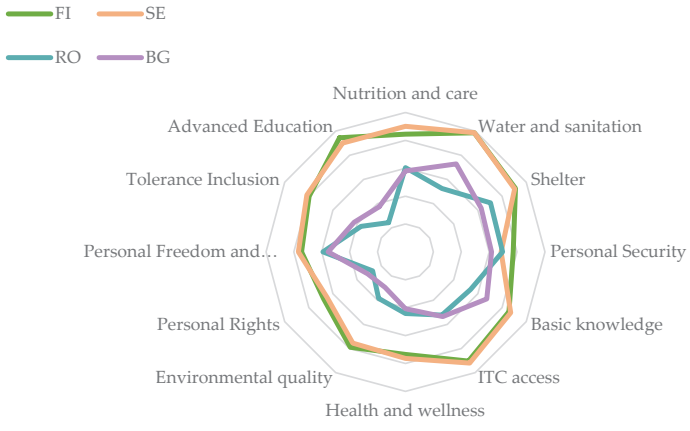


Figure 2. Social Progress Index for 2020.

Regarding the evolution of the SPI for the analyzed period, some improvements can be observed for the least-performing countries analyzed in terms of access to basic knowledge, information and communications. Therefore, we can admit, ceteris paribus, that the absorption of cohesion funds positively influenced social progress, as measured by the SPI.

According to the latest SPI global data (Figure 3), among the countries of Central and Eastern Europe, the best-performing country is Estonia, which ranks 21st globally, followed by the Czech Republic—22nd, Slovenia—26th, Lithuania—27th, Slovakia—33rd, Latvia—34th, Poland—35th, Croatia—36th, Hungary—42nd, Bulgaria—43rd, and Romania—44th.

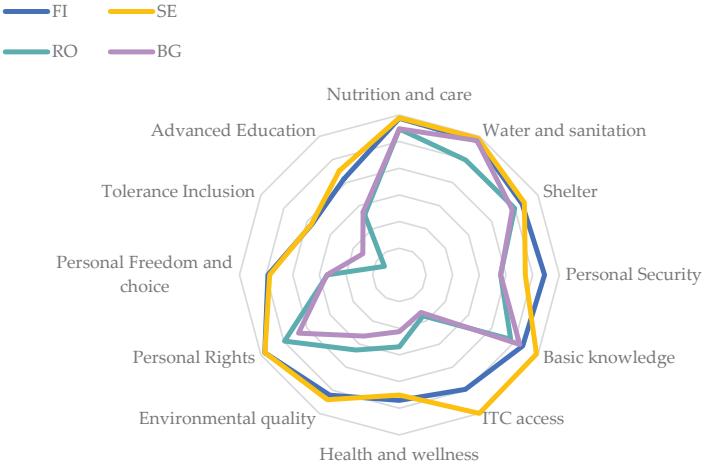


Figure 3. Social Progress Index—2014.

The EQI, as delineated by the European Commission, reflects citizens' collective perceptions and encounters concerning corruption and the quality and fairness of pivotal public services—namely health, education, and law enforcement—within their residing regions. This index is constructed from the most expansive survey ever conducted to gauge the quality perception of governance within the EU. This monumental survey amassed insights and firsthand experiences in public health, education, and law enforcement from over 129,000 participants across 208 regions spanning all 27 EU Member States, assessed at either the NUTS1 or NUTS2 level. The survey's inquiries are rooted in a comprehensive, multifaceted understanding of government quality, encompassing high standards of impartiality and public service delivery paired with minimal corruption.

Figure 4 provides an overview of the quality of public services in the EU regions, highlighting a segregation between South-East and North-West Europe that can be associated with the level of development of the regions as the South-East area is characterized by transition countries that are less developed. In contrast, the North-West area represents the hard core of the EU. The EQI level is of considerable importance in the context of the absorption and use of European funds, as a lower value of government quality can undermine the efficiency of the use of financial resources or even generate a deviation from the cohesion objective. The lower EQI values observed in the Southeastern regions, associated with transition countries, are supported by various studies, including those by Acemoglu and Robinson (2012), who emphasize the impact of governance quality on economic outcomes. Additionally, works by Zaman and Georgescu (2009), Crescenzi and Giua (2016), and Calegari (2020) reinforce the idea that government quality is instrumental in determining the success of regional development policies.

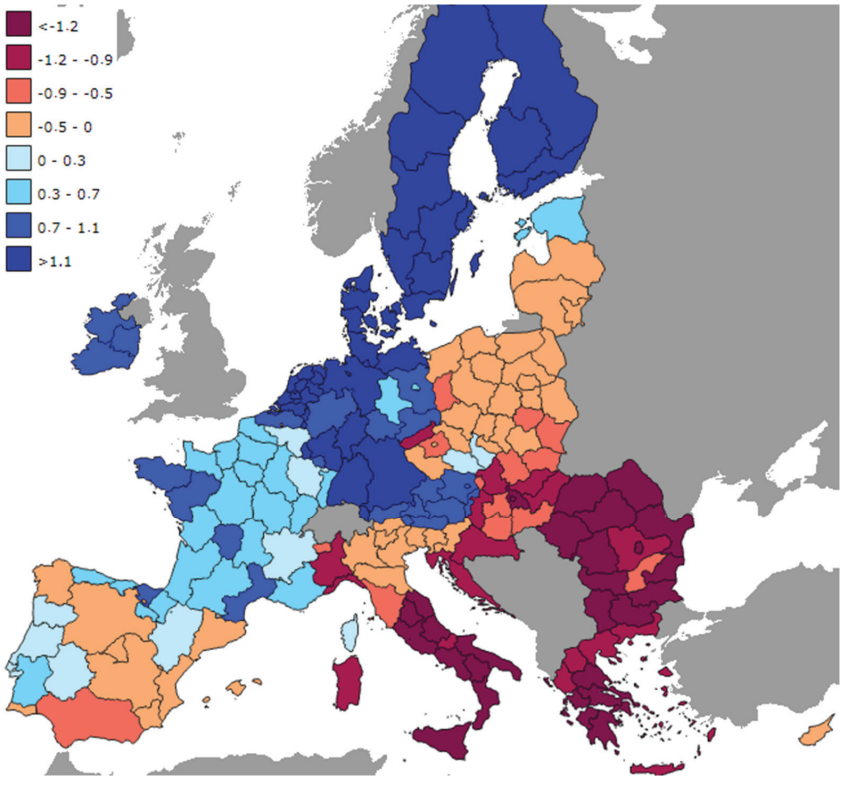


Figure 4. European Government Quality Index (EQI 2017).

Furthermore, the discussion surrounding the segregation between South-East and North-West Europe is rooted in the historical context of EU expansion, with the Southeastern regions often facing challenges in catching up with their more developed counterparts. Empirical data from various studies, such as those by Rodríguez-Pose and Hardy (2015) and Becker et al. (2013), support the notion that the disparities in government quality contribute to variations in the absorption and effective utilization of European funds.

The future direction of the cohesion policy in 2021–2027 will focus mainly on the green and digital transition. The 2021–2027 multiannual financial framework was geared towards rapid recovery and resilience, strengthening convergence between EU regions, given the uneven territorial impact of the crisis. In this context, the legislative package on cohesion policy for the 2021–2027 programming period entered into force on July 1, 2021. The newly reformed rules were designed to increase the focus of cohesion policy on a "smarter" and "greener" Europe and to create favorable conditions for investment with simplified delivery mechanisms and closer links to structural reforms. Cohesion policy will thus contribute to implementing the EU political agenda, particularly promoting the green transition and digital transformation. The EU has committed to becoming the world's first climate-neutral bloc by 2050. A Just Transition Fund has been set up under cohesion policy to ensure that the transition to a climate-neutral economy equitably takes place, leaving no region behind.

4.2. Evaluation of the Impact of Cohesion Policy Instruments on the Economic and Social Performance of CEE Countries

Different specifications have been tested in terms of components and total value to assess the impact of cohesion policy on economic growth and social progress. In every one of the ten models evaluated, we employed the Hausman test to determine the nature of the effects (be it fixed or random). The outcomes consistently pointed to the existence of random effects across all models. Additionally, even though the models were deemed valid, adjustments were made to account for heteroscedasticity.

In the case of economic growth, the empirical results highlighted in all four analyzed models the positive and significant impact of the investments of the funds specific to the cohesion policy (ERDF, CF and ESF) on economic growth, both in terms of components and in total value (Appendix B, Table A2).

Thus, if we refer to the total investment of cohesion policy funds per inhabitant, an increase in this indicator by 1% leads to an increase in economic performance, ceteris paribus, by 0.05%. CF investments have the most pronounced impact in terms of investment type, they lead to an increase in economic performance, ceteris paribus, by 0.04%.

The positive coefficients associated with the specific funds highlight the importance of European funds in driving regional economic growth. This implies that the strategic allocation of these funds can foster positive economic outcomes for regions that receive them. Each fund had a distinct impact on economic growth, as evidenced by the varying coefficients across models. This suggests that while all funds contribute positively, they have different areas of focus or effectiveness. While the ERDF primarily focuses on economic and social cohesion by correcting regional imbalances, the CF targets environmental and trans-European transport infrastructure in countries with a lower GDP.

Given that specific funds like the ERDF support regions with natural handicaps or those that are demographically disadvantaged, the positive impact on real GDP per capita underscores the importance of such funds in bridging regional disparities and ensuring more equitable growth across regions.

The positive impact of these funds on GDP per capita might reinforce the need for policymakers to continue prioritizing them. Moreover, understanding which funds have the most significant impact can guide future budgetary allocations. While the funds have a positive impact, the magnitude of the impact depends on how efficiently these funds are utilized. Efficient project implementation and fund utilization can maximize the economic benefits.

In essence, the observed positive impacts of specific European funds on real GDP per capita stress the pivotal role these funds play in regional economic development. It emphasizes the need for continued investment, effective utilization, and strategic allocation to maximize the desired economic outcomes.

The initial GDP per capita (from the previous period) has a positive and statistically significant impact on the current real GDP per capita. A 1% increase in the initial GDP per capita leads to increases in the current real GDP per capita, ranging from 0.34% to 0.4% across models. This suggests that regions with a higher initial GDP maintain economic strength over time.

The EQI has a consistently positive and statistically significant effect on the Real GDP per capita across all models. The influence of EQI ranges from 0.13% to 0.18% for a unit increase, highlighting the importance of governance quality in positively influencing economic performance. Higher values of the EQI, indicating better governance and public service delivery, are associated with higher real GDP per capita.

In summary, as measured by the EQI, good governance quality and investments from various European funds, especially when initiated in the previous period, have a positive and significant impact on the economic performance of regions. This highlights the importance of effective governance and strategic investment in economic growth.

In terms of capturing the impact of the ESIF (European Structural and Investment Fund) investments on social progress, the empirical results revealed a positive and statistically significant impact that was preserved in all four models (both in total and by components of ESIF investments). The greatest impact on social progress was manifested by CF investment, ceteris paribus. The results indicate that the investments made by the ESIF positively influence social progress in the regions. This means that as the funding from the ESIF increases or is efficiently utilized, there is a consequent improvement in social indicators like education, health, equality, and overall well-being. The statement emphasizes that this positive effect is consistent across all four models they analyzed. Whether looking at the total investments or breaking them down into specific components, the positive relationship holds. This consistency reinforces the reliability of the findings. Among the ESIF components, the CF has the most pronounced positive effect on social progress. This might be because the CF specifically targets areas like environmental projects and trans-European transport networks, which can have direct or trickle-down effects on the populace's well-being.

Also, the initial value of real GDP/capita and the EQI value showed the same positive and statistically significant impact. So, the regions that registered a good economic start and benefited from high-quality institutions achieved increased social progress.

Regions that began the period with a higher GDP per capita (a general measure of economic well-being) also saw greater social progress. This suggests a positive feedback loop where regions with a good economic base can better leverage ESIF investments to further social advancement.

The quality of institutions, as measured by the EQI, also played a significant role. Regions with better governance, lower corruption, and higher quality public services (like health, education, and law enforcement) were more successful in translating ESIF investments into social progress. Good governance can ensure the efficient and effective use of funds, leading to tangible improvements in the lives of the citizens.

The findings underscore the importance of financial investments from funds like the ESIF and foundational factors like initial economic conditions and institutional quality in driving regional social progress.

This research conducted a thorough robustness assessment, further probing the effects of ESIF investments on the various facets of the social progress indicator (as seen in Table A3, Appendix C). The empirical results demonstrating a positive and significant influence on Basic Human Needs and Fundamental Well-being underscore the effectiveness of cohesion policies in enhancing aspects such as healthcare, education, and environmental quality, which are essential for the foundational welfare of individuals in CEE nations. Also,

this positive impact reflects the direct benefits individuals in CEE regions experience in their daily lives, such as better access to basic services and improved living conditions, validating the targeted approach of cohesion policies towards addressing fundamental human necessities.

However, the lack of a significant impact on the Opportunities segment of the SPI, particularly in areas related to access to higher education and information technology, points to a potential disconnect between the economic growth facilitated by cohesion policies and the translation of this growth into real opportunities for individuals, especially in terms of advancing equity and reducing income disparities.

The consistently lower scoring of CEE nations in the SPI's Opportunities dimension highlights a critical area for policy improvement, suggesting that while basic needs and well-being are being addressed, the policies are less effective in creating environments where individuals can capitalize on economic growth to improve their personal and professional prospects. This gap in policy impact, especially in the realm of reducing income inequality and tackling poverty, calls for a re-evaluation of the current policy frameworks and strategies, emphasizing the need to not only foster economic growth but to also ensure that such growth translates into equitable opportunities and tangible benefits for all segments of society, particularly the most vulnerable.

This observation echoes Atkinson's assertion (Atkinson 2015) that disparities in outcomes observed today pave the way for unequal opportunities in the future. This concept is particularly relevant in the context of income inequality and poverty. When current disparities in income and wealth are not addressed, they perpetuate a cycle of poverty and limit opportunities for upward social mobility for future generations. Therefore, the focus on enhancing interventions in the Opportunities sector of the SPI is crucial. By actively reducing income disparities and improving access to education, technology, and fair employment, we can create a more equitable foundation that supports the well-being and socio-economic advancement of upcoming generations. This approach is not just about reducing present inequalities; it is a strategic investment in preventing the entrenchment of poverty and ensuring fairer, more equitable opportunities for all, thereby breaking the cycle of poverty and income inequality that can otherwise persist across generations.

5. Discussions and Policy Implications

While evaluating the effects of cohesion policy is complex due to its multidimensional nature, targeting both economic and social goals, it is crucial to note that these varied interventions have historically emphasized economic development over direct social impacts, particularly in addressing income inequality and poverty. The challenge lies in distinguishing the nuanced effects on social outcomes, such as reducing income disparities and enhancing social welfare, often overshadowed by the primary focus on economic indicators like GDP.

The academic discourse, including works by (Zaman and Georgescu 2009; Antonescu 2012; Crescenzi and Giua 2016; Calegari 2020; Maucorps et al. 2020), predominantly examines the economic impacts of EU cohesion policy. However, this focus leaves an empirical gap regarding the policy's effectiveness in combating social issues like poverty and inequality. While some studies (Pinho et al. 2015; Crescenzi and Giua 2016; Cerqua and Pellegrini 2018; Fidrmuc et al. 2019; Di Caro and Fratesi 2022) suggest a positive economic impact, the direct correlation with social improvements, particularly in reducing poverty and narrowing income gaps, remains underexplored (Dall'Erba and Fang 2017; Ehrlich and Overman 2020).

Our study's alignment with the literature, such as Farrell (2004) and Lolos (2009), confirming positive economic impacts, also highlights a need to deepen the understanding of how these economic benefits translate into tangible social progress. While economic development is a positive outcome, its success should also be measured by its ability to alleviate poverty and improve income equality. As with the study by Maucorps et al. (2020), our results, which took into account the impact of the three funds specific to the cohesion

policy on economic development reflected through GDP per capita at the regional level (NUTS2) (elements that were taken into account in both studies), identify a positive impact.

Regarding the impact of cohesion policy on social progress, our study identified a positive impact on the aggregate SPI and the two pillars of the SPI (Basic Human Needs and Basic Well-being). Still, for the countries considered (CEE), an impact on the Opportunities pillar was not identified. This gap is particularly critical when considering income inequality and poverty, as the Opportunities pillar directly relates to aspects like access to higher education and personal rights, which are key in breaking the cycle of poverty and ensuring equitable growth.

It is difficult to compare the results with other studies as the literature relating to the impact on the SPI is limited. However, some studies have considered some of the variables included in the SPI. Most of the literature related to the impact of European funds on the employment rate (Becker et al. 2010; Rodríguez Pose and Novak 2013; Giua 2017; Fiaschi et al. 2018; Crescenzi and Giua 2018; Crescenzi and Giua 2020) predominantly identify a positive impact, considering that European funds contribute to increasing the employment rate. Other social aspects that the researchers focused on were the level of education and the modified version of the adapted human development index (Calegari 2020), where a positive impact was also identified.

Therefore, the consistent positive impact associated with cohesion policy on economic growth emphasizes the crucial role of this policy in fostering regional development. Both economic growth and social progress have benefited from the funds, implying that well-structured policy interventions can produce positive outcomes in these areas.

Regarding the impact on economic growth and social outcomes, this study confirms a positive correlation between the EU's cohesion policy and economic growth in CEE countries, aligning with the existing literature, such as Farrell (2004) and Lolos (2009). However, when it comes to the policy's impact on social outcomes, particularly in reducing poverty and income disparity, the findings reveal a more complex scenario. While there is a positive impact on the aggregate Social Progress Index (SPI) and its components of Basic Human Needs and Well-being, the gap in the Opportunities pillar suggests that the policy's effectiveness in addressing deeper social inequities needs more focus.

In terms of the correlation between policy implementation and regional improvements, the empirical results indicate that regions with a stronger economic foundation (higher initial GDP per capita) tend to leverage the benefits of cohesion funds more effectively, thereby amplifying their economic and social gains. This finding underscores the need for quality governance and efficient institutions, which are essential for optimizing fund utilization.

Concerning the time-lagged nature of policy impact, this study suggests that the impact of cohesion policy on social progress, especially in terms of improving personal rights and access to higher education, is not immediate but evolves over time. This highlights the need for a long-term strategic vision in policy implementation and assessment.

The CF significantly impacts economic performance among the funds examined. This underscores the CF's pivotal role in addressing disparities and promoting convergence among EU regions, especially in environmental projects and trans-European transport networks.

Regions with a good economic start (higher initial GDP per capita) tend to maintain and amplify their economic strengths. Further, the consistent positive impact of the EQI underscores the importance of quality governance in driving economic outcomes. Efficient institutions that maintain low corruption levels and provide high-quality public services pave the way for the better utilization of funds and stronger economic growth.

The results indicate that regions effectively utilizing ESIF investments witness marked social progress. A holistic approach, combining financial investments from funds like ESIF with a solid initial economic base and high institutional quality, can catalyze substantial social advancements.

The research findings carry substantial significance for shaping the trajectories of social and economic progress in CEE economies. By confirming the consistent positive impact of EU cohesion policy on economic growth in the region, this study underscores

the pivotal role of targeted interventions and continued support in fostering regional development. This insight is particularly crucial for policymakers as it provides evidence of the effectiveness of cohesion policy in narrowing economic disparities and promoting overall economic advancement in CEE countries.

The research shows a gap in the Opportunities segment of the social progress indicator for CEE countries. This suggests a need for targeted interventions in these regions, particularly focusing on personal rights, electoral choices, tolerance, and access to higher education. As emphasized by Atkinson's observation (Atkinson 2015), addressing today's inequalities is crucial for ensuring equal opportunities for future generations.

Taking into account the results of the study, we consider it important that, for CEE countries and regions, the following aspects are taken into account in the management of the cohesion policy:

- Increase Absorption of European Funds for Economic and Social Welfare: This research emphasizes the need for greater efforts to increase the absorption degree of European funds. This recommendation aligns with the findings that regions effectively utilizing ESIF allocations witness marked social progress. The suggestion to augment absorption aims to leverage these funds more effectively to contribute to both economic development and social welfare.
- Enhance Institutional Quality and Transparency: The study acknowledges the positive impact of efficient institutions on economic outcomes. Therefore, the recommendation to focus on measures and reforms to enhance institutional quality aligns with the research findings. It emphasizes the importance of quality governance in driving economic growth and ensuring better utilization of funds.
- Streamline Bureaucratic Procedures: The suggestion to streamline public administration processes and reduce bureaucratic barriers resonates with the research findings calling for de-bureaucratization. Simplifying access to European funds can facilitate their efficient utilization, which, as the research indicates, contributes to regional economic and social progress.
- Concentrate Investments in Social Areas with Poor Infrastructure: The research highlights a gap in the Opportunities segment of the SPI for CEE countries, particularly in areas like access to higher education and personal rights. The recommendation to concentrate investments in social areas with poor infrastructure aligns with this finding, aiming to reduce regional economic and social discrepancies and ultimately promote increased cohesion.
- Targeted Interventions to Address Inequalities: The research underscores the need for targeted interventions in specific aspects such as personal rights, tolerance, and access to higher education. The policy recommendation advocating for a more nuanced approach to cohesion policy in CEE countries aligns with the findings, emphasizing the necessity of targeted interventions to directly address social disparities and ensure equal opportunities.

Given these findings, our research advocates for a more nuanced approach to cohesion policy in CEE countries, emphasizing the need for targeted interventions that boost economic development and directly address social disparities. This includes improving institutional quality and transparency, reducing bureaucratic barriers to fund access, and prioritizing investments in social infrastructure. Such a comprehensive strategy is essential for bridging economic and social gaps, leading to a more cohesive society where economic growth translates into reduced poverty and greater income equality.

6. General Conclusions

The empirical analysis supports the notion that well-structured EU cohesion policy interventions can produce positive outcomes in both economic and social domains. The significant impact of funds like the European Regional Development Fund (ERDF) on regional economic performance points to their crucial role in promoting convergence and addressing disparities, particularly in environmental and transport projects.

From this study's analysis, it is evident that cohesion policy positively influences both economic growth and social progress. However, it is crucial to acknowledge that the impacts of this policy are gradual, often manifesting over several years. Investments like creating new job opportunities take time to generate economic benefits, and their effectiveness is closely tied to addressing income inequality and poverty reduction.

To maximize the benefits of cohesion policy, policymakers should consider strategic allocations, emphasizing funds that have shown pronounced impacts, like the CF, which has shown considerable potential in promoting regional development, including reducing the risk of poverty.

However, this study also identifies a notable gap in the Opportunities segment of the SPI for CEE countries. This gap is critical as it directly relates to aspects such as access to higher education and personal rights, which are vital for breaking the cycle of poverty and ensuring equitable growth. Addressing this gap requires targeted interventions focusing on enhancing personal rights, electoral choices, tolerance, and access to higher education.

Further, addressing the gap in the Opportunities dimension within CEE countries is essential. This is not just about economic development but also about ensuring equitable access to opportunities, which is crucial for long-term poverty alleviation and narrowing income gaps. The findings highlight the need for a cohesive approach, combining strategic financial investments with solid economic fundamentals and effective governance, to ensure regional development, social equity, and reduced income disparities.

Regarding research limitations, one of the foremost challenges faced during the research was the accessibility and completeness of data. Both data sets related to the cohesion funds and social indicators were limited. Also, the limited number of existing studies exploring the social impact of cohesion funds is a significant constraint that curtails the ability to juxtapose current research findings with previous work. The time it takes for the effects of the cohesion funds to become visible can be prolonged. Determining this lag between investment and observable impact can be complex, especially when it intertwines with other external factors. This makes attributing specific outcomes to cohesion funds difficult without considering potential interactions with other initiatives.

The research findings have substantial implications for policymakers, highlighting the effectiveness of cohesion policy in narrowing economic disparities and promoting overall economic advancement in CEE countries. However, they also call for an enhanced focus on social dimensions, particularly in areas where progress has been limited.

In conclusion, our study reinforces the importance of a comprehensive approach to assessing cohesion policy, one that goes beyond traditional economic metrics to include a broader spectrum of social indicators. This approach is vital for accurately gauging the policy's impact and for informing future policy decisions aimed at promoting balanced and inclusive growth across the CEE region.

As future directions of research, given the European Commission's efforts to simplify the cohesion policy, it would be valuable to investigate the impacts of such simplifications. Specifically, examining the consequences of reducing the number of objectives and focusing on digitization and greening can provide insights into the policy's efficacy under this streamlined approach.

Investigating the relationship between administrative capacity and the efficiency of fund utilization is also crucial, particularly in how it relates to the equitable distribution of economic gains and poverty reduction.

Given that the quality of institutions plays a pivotal role in the effective use of financial resources and in addressing inequalities, future research can explore strategies to bolster institutional quality.

Another research direction, given the importance of the Opportunities pillar in reducing income inequalities, is for researchers to focus on understanding its components better. For instance, which aspects of "Opportunities" (like access to advanced education or respect for personal rights) have the most pronounced effect on income inequalities? Such insights can guide targeted interventions.

This study has hinted at the potential of cohesion policies to address income inequalities for future generations. A logical extension would be to analyze the long-term impact of these policies on income distribution.

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

Funding: The research study has been elaborated within the Data Science Research Lab for Business and Economics of the Bucharest University of Economic Studies, within project ID 585 PERFECTIS, entitled 'Increasing Institutional Performance through the Development of the Infrastructure and Research Ecosystem of Transdisciplinary Excellence in the Socio-economic Field', Funder: Ministry of Research, Innovation and Digitization, contract number and date: 42PFE of 30 December 2021.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data are contained within the article.

Conflicts of Interest: The authors declare no conflicts of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

Abbreviations

Abbreviations	Descriptions
EU	European Union
GDP	Gross Domestic Product
CEE	Central and Eastern European
ERDF	European Regional Development Fund
SPI	Social Progress Index
PIGS	Portugal, Ireland, Greece and Spain
HDI	Human Development Index
CF	Cohesion Fund
ESF	European Social Fund
ESIF	European Structural and Investment Funds
EQI	Governance Quality Index
GNI	Gross National Income

Appendix A

Table A1. Description of the variable.

Indicator	Definition	Source	Expected Impact
	Dependent variables		
	Economic Performance		
Real GDP/cap.	Real gross domestic product expressed according to no. inhabitants	Eurostat	(+)
	Social Progress		
Social Progress Index (SPI)	The Social Progress Index measures the progress of society in three dimensions (Basic Human Needs, Foundations of Well-being and Opportunities)	Social Progress Imperative	(+)

Table A1. *Cont.*

Indicator	Definition	Source	Expected Impact
Basic Human Needs	The "Basic Human Needs" dimension of SPI	Social Progress Imperative	(+)
Foundations of Well-being	The "Foundations of Well-being" dimension of SPI	Social Progress Imperative	(+)
Opportunity	The "Opportunity" dimension of SPI	Social Progress Imperative	(+)
EQI	European Governance Quality Index	GoQ Institute	Control variable
	Independent variables		
	Cohesion Policy		
ERDF	Amount of allocations from the European Regional Development Fund (euro—deflated values)	European Commission	
CF	Amount of allocations from the Cohesion Fund (euro—deflated values)	European Commission	
ESF	The amount of allocations from the European Social Fund (euro—deflated values)	European Commission	

Appendix B

Table A2. Empirical results of the impact of cohesion policy on economic performance and social progress.

Variable	M1	M2	M3	M4	M5	M6	M7	M8
	Dep. Var: Log(Real GDP/cap.) t				Var. Dep. Social Progress Index (SPI)			
EQI t	0.15 ***	0.17 ***	0.13 ***	0.18 ***	1.21 ***	1.09 ***	1.20 ***	0.91 ***
Log (Int.GDP/cap.) t − 1	0.4 **	0.34 **	0.34 **	0.35 **	3.6 ***	3.58 ***	3.6 ***	3.58 ***
Log (inv. Tot.ESIF/cap.) t − 1	0.05 ***					0.25 ***		
Log (ERDF/cap.) t − 1		0.03 ***					0.13 ***	
Log (CF/cap.) t − 1			0.04 ****					0.31 ***
Log (FES/cap.) t − 1				0.013 ***	0.10 **			
No. obs.	594	592	585	590	590	594	592	585
No. country	10	10	10	10	10	10	10	10
No. Reg.	54	54	54	54	54	54	54	54
No. year	12	12	12	12	12	12	12	12
Wald Test	95.41 ***	78.37 ***	88.36 ***	51.30 ***	72.46 ***	162.84 ***	97.60 ***	220.83
R^2								
-within	0.2276	0.1854	0.2405	0.1492	0.1593	0.1952	0.1648	0.2616
-between	0.3123	0.3173	0.3204	0.2988	0.3853	0.3856	0.3828	0.3940
-overall	0.3082	0.3154	0.3079	0.2951	0.3738	0.3760	0.3739	0.3800
F Test	208.67 ***	195.76 ***	199.65 ***	183.36 ***	213.56 ***	221.06	215.07 ***	230.80 ***
Hausman Test prob.	0.473	0.411	0.472	0.267	0.28	0.244	0.223	0.380
Breusch–Pagan LM Test pt. RE	2617.77 ***	2525.61 ***	2527.66 ***	2296.82 ***	2482.95 ***	2514.59 ***	2468.99 ***	2493.58 ***
Cross-sectional Dependence Test—Pesaran CD Test	80.005 ***	81.48 ***	70.85 ***	81.46 ***	78.20 ***	70.09 ***	74.23 ***	62.85 ***
Modified Wald Test for Groupwise Heteroskedasticity in Fixed-Effect Regression Model	14,626.43 ***	12,297 ***	8329.32 ***	12,403.86 ***	13,738.81 ***	16,802.97	14,651.51 *** 0.0000	10,265.35 ***

Source: Own calculations; Note: ****, *** and ** means significant at 0.5%, 1% and 5% level; () represents the probability.

Appendix C

Table A3. Empirical results of the impact of cohesion policy on social progress (robustness analysis).

Variable	M1 Basic Human Needs	M2 Foundations of Well-Being
EQI t	0.58 ***	2.99 ***
Log (Int.GDP/cap.) t	1.87 ***	4.31 ***
Log (inv. Tot.FESI/cap.) t − 1	0.23 ***	0.59 ***
No. obs.	594	594
No. country	10	10
No. Reg.	54	54
Nr. ani	12	12
Wald Test PROB.	110.40 ***	188.06 ***
R^2		
-within	0.1521	0.2085
-between	0.2176	0.4860
-overall	0.2124	0.4276
F Test	147.99 ***	41.75 ***
Hausman Test prob.	0.11	0.31
Breusch–Pagan LM Test pt. RE	2382.78 ***	1261.66 ***
Cross-sectional Dependence Test—Pesaran CD Test	80.86 ***	93.25 ***
Modified Wald Test for Groupwise Heteroskedasticity in Fixed-Effect Regression Model	30,454.34 ***	2495.04 ***

Source: Own calculations; Note: *** means significant at 1% level; () represents the probability.

Notes

1 https://en.wikipedia.org/wiki/Human_Development_Index. Accessed on 12 September 2023.

2 https://hdr.undp.org/data-center/human-development-index#/indicies/HDI. Accessed on 14 September 2023.

3 "Beyond GDP". The Economist. 18 April 2013. Retrieved 2 August 2013.

4 https://www.socialprogress.org/. Accessed on 6 September 2023.

References

Acemoglu, Daron, and James A. Robinson. 2012. Why nations fail: The origins of power, prosperity, and poverty. *Finance and Development-English Edition* 49: 53. [CrossRef]

Albanese, Giuseppe, Vincenzo Carrieri, and Maria Speziali. 2021. Looking for a Star: Evaluating the Effect of the Cohesion Policy on Regional Well-Being. Available online: https://www.econstor.eu/handle/10419/245572 (accessed on 6 September 2023).

Albulescu, Claudiu, and Daniel Goyeau. 2013. EU funds absorption rate and the economic growth. *Timisoara Journal of Economics and Business* 6: 153–70.

Antonescu, Daniela. 2012. Identifying regional disparities in Romania: A convergence process perspective in relation to European Union's territorial structures. *Procedia Economics and Finance* 3: 1148–55. [CrossRef]

Atkinson, Anthony B. 2015. *Inequality: What Can Be Done?* Cambridge: Harvard University Press.

Barca, Fabrizio. 2009. *Agenda for a Reformed Cohesion Policy*. Brussels: European Communities. Available online: https://migrant-integration.ec.europa.eu/library-document/agenda-reformed-cohesion-policy-place-based-approach-meeting-european-union_en (accessed on 10 September 2023).

Becker, Sascha O., Peter H. Egger, and Maximilian Von Ehrlich. 2010. Going NUTS: The effect of EU Structural Funds on regional performance. *Journal of Public Economics* 94: 578–90. [CrossRef]

Becker, Sascha O., Peter H. Egger, and Maximilian Von Ehrlich. 2013. Absorptive capacity and the growth and investment effects of regional transfers: A regression discontinuity design with heterogeneous treatment effects. *American Economic Journal: Economic Policy* 5: 29–77. [CrossRef]

Boldrin, Michele, and Fabio Canova. 2001. Inequality and convergence in Europe's regions: Reconsidering European regional policies. *Economic Policy* 16: 206–53. [CrossRef]

Calegari, Elena. 2020. *Regional Policies and Well-Being: An Evaluation Approach*. Piacenza: Università Cattolica del Sacro Cuore.

Calegari, Elena, Enrico Fabrizi, Gianni Guastella, and Francesco Timpano. 2021. EU regional convergence in the agricultural sector: Are there synergies between agricultural and regional policies? *Papers in Regional Science* 100: 23–50. [CrossRef]

Cerqua, Augusto, and Guido Pellegrini. 2018. Are we spending too much to grow? The case of Structural Funds. *Journal of Regional Science* 58: 535–63. [CrossRef]

Charron, Nicholas. 2016. Explaining the allocation of regional Structural Funds: The conditional effect of governance and self-rule. *European Union Politics* 17: 638–59. [CrossRef]

Charron, Nicholas, Lewis Dijkstra, and Victor Lapuent. 2014. Regional governance matters: Quality of government within European Union member states. *Regional Studies* 48: 68–90. [CrossRef]

Crescenzi, Riccardo, and Andrés Rodríguez-Pose. 2012. Infrastructure and regional growth in the European Union. *Papers in Regional Science* 91: 487–513. [CrossRef]

Crescenzi, Riccardo, and Mara Giua. 2016. The EU Cohesion Policy in context: Does a bottom-up approach work in all regions? *Environment and Planning A: Economy and Space* 48: 2340–57. [CrossRef]

Crescenzi, Riccardo, and Mara Giua. 2018. *One or Many Cohesion Policies of the European Union? On the Diverging Impacts of Cohesion Policy across Member States*. London: Centre for Economic Performance, LSE.

Crescenzi, Riccardo, and Mara Giua. 2020. One or many Cohesion Policies of the European Union? On the differential economic impacts of Cohesion Policy across member states. *Regional Studies* 54: 10–20. [CrossRef]

Dall'Erba, Sandy, and Fang Fang. 2017. Meta-analysis of the impact of European Union Structural Funds on regional growth. *Regional Studies* 51: 822–32. [CrossRef]

Dall'Erba, Sandy, and Julie Le Gallo. 2008. Regional convergence and the impact of European structural funds over 1989–1999: A spatial econometric analysis. *Papers in Regional Science* 87: 219–44. [CrossRef]

Darvas, Zsolt, and Guntram B. Wolff. 2018. The EU's Multiannual Financial Framework and Some Implications for CESEE Countries. Focus on. Available online: https://ideas.repec.org/a/onb/oenbfi/y2018iq3-18b7.html (accessed on 6 September 2023).

Di Caro, Paolo, and Ugo Fratesi. 2022. One policy, different effects: Estimating the region-specific impacts of EU cohesion policy. *Journal of Regional Science* 62: 307–30. [CrossRef]

Dinu, Marin, Cristian Socol, and Marius Marinaş. 2005. *Mecanisme de convergenţă şi coeziune*. Bucharest: Economica.

Ederveen, Sjef, Henri L. F. De Groot, and Richard Nahuis. 2006. Fertile soil for structural funds? A panel data analysis of the conditional effectiveness of European cohesion policy. *Kyklos* 59: 17–42. [CrossRef]

Ederveen, Sjef, Joeri Gorter, Ruud De Mooij, and Richard Nahuis. 2003. Funds and Games. The Economics of European Cohesion Policy. Available online: https://www.files.ethz.ch/isn/111108/003_Funds%20and%20Games.pdf (accessed on 6 September 2023).

Ehrlich, Maximilian V., and Henry G. Overman. 2020. Place-based policies and spatial disparities across European cities. *Journal of Economic Perspectives* 34: 128–49. [CrossRef]

Farrell, Mary. 2004. Regional integration and cohesion—Lessons from Spain and Ireland in the EU. *Journal of Asian Economics* 14: 927–46. [CrossRef]

Ferrara, Antonella Rita, Lewis Dijkstra, Philip McCann, and Rosanna Nisticó. 2022. The response of regional well-being to place-based policy interventions. *Regional Science and Urban Economics* 97: 103830. [CrossRef]

Fiaschi, Davide, Andrea Mario Lavezzi, and Angela Parenti. 2018. Does EU cohesion policy work? Theory and evidence. *Journal of Regional Science* 58: 386–423. [CrossRef]

Fidrmuc, Jan, Martin Hulényi, and Olga Zajkowska. 2019. *The Elusive Quest for the Holy Grail of an Impact of EU Funds on Regional Growth*. CESifo Working Paper No. 7989. Munich: Center for Economic Studies and IFO Institute.

Fratesi, Ugo, and Fiona G. Wishlade. 2017. The impact of European Cohesion Policy in different contexts. *Regional Studies* 51: 817–21. [CrossRef]

Giua, Mara. 2017. Spatial discontinuity for the impact assessment of the EU regional policy: The case of Italian objective 1 regions. *Journal of Regional Science* 57: 109–31. [CrossRef]

Klugman, J. 2011. Human Development Report 2011. Sustainability and Equity: A better future for all. Sustainability and Equity: A Better Future for All. *UNDP-HDRO Human Development Reports*. Available online: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2294671 (accessed on 10 September 2023).

Lolos, Sarantis E. G. 2009. The effect of EU structural funds on regional growth: Assessing the evidence from Greece, 1990–2005. *Economic Change and Restructuring* 42: 211–28. [CrossRef]

Manzella, Gian Paolo, and Carlos Mendez. 2009. The Turning Points of EU Cohesion Policy. Working Paper Report to Barca Report. Available online: https://strathprints.strath.ac.uk/38743/ (accessed on 12 September 2023).

Maucorps, Ambre, Stefan Jestl, and Roman Römisch. 2020. *The Effects of the EU Cohesion Policy on Regional Economic Growth: Using Structural Equation Modelling for Impact Assessment*. wiiw Working Paper. Vienna: The Vienna Institute for International Economic Studies, wiiw.

Medeiros, Eduardo. 2014. Assessing territorial impacts of the EU Cohesion Policy: The Portuguese case. *European Planning Studies* 22: 1960–88. [CrossRef]
Nistotskaya, Marina, Nicholas Charron, and Victor Lapuente. 2015. The wealth of regions: Quality of government and SMEs in 172 European regions. *Environment and Planning C: Government and Policy* 33: 1125–55. [CrossRef]
Piketty, Thomas. 2014. *Capital in the Twenty-First Century*. Cambridge: Harvard University Press.
Pinho, Carlos, Celeste Varum, and Micaela Antunes. 2015. Under what conditions do structural funds play a significant role in European regional economic growth? Some evidence from recent panel data. *Journal of Economic Issues* 49: 749–71. [CrossRef]
Pose, Andrés Rodríguez, and Katja Novak. 2013. Learning processes and economic returns in European Cohesion policy. *Investigaciones Regionales = Journal of Regional Research* 25: 7–26.
Rodríguez-Pose, Andrés, and Daniel Hardy. 2015. Addressing poverty and inequality in the rural economy from a global perspective. *Applied Geography* 61: 11–23. [CrossRef]
Sachs, Jeffrey D. 2015. *The Age of Sustainable Development*. New York: Columbia University Press.
Sapir, André. 2003. *Report of an Independent High-Level Study Group Establish on the Initiative of the President of the European Commission*. Available online: https://www.researchgate.net/publication/227467942_An_Agenda_for_a_Growing_Europe_The_Sapir_Report (accessed on 15 September 2023).
Sen, Amartya. 1999. Development as freedom. *The Globalization and Development Reader: Perspectives on Development and Global Change* 2014: 525.
Stiglitz, Joseph E., Amartya Sen, and Jean-Paul Fitoussi. 2010. *Mismeasuring Our Lives: Why GDP Doesn't Add Up*. New York: The New Press.
Surubaru, Neculai-Cristian. 2017. Administrative capacity or quality of political governance? EU Cohesion Policy in the new Europe, 2007–13. *Regional Studies* 51: 844–56. [CrossRef]
Tomova, Mariana, András Rezessy, Artur Lenkowski, and Emmanuelle Maincent. 2013. *EU Governance and EU Funds-Testing the Effectiveness of EU Funds in a Sound Macroeconomic Framework*. Bruxelles: Directorate General Economic and Financial Affairs (DG ECFIN), European Commission.
Zaman, Gheorghe, and George Georgescu. 2009. Structural fund absorption: A new challenge for Romania. *Romanian Journal of Economic Forecasting* 1: 136–54.

economies

MDPI

Article

Does Gender Equality in Managerial Positions Improve the Gender Wage Gap? Comparative Evidence from Europe

Lidia de Castro Romero *, Víctor Martín Barroso and Rosa Santero-Sánchez

Department of Applied Economic I, Faculty of Economics and Business Sciences, Rey Juan Carlos University, 28032 Madrid, Spain; victor.martin@urjc.es (V.M.B.); rosa.santero@urjc.es (R.S.-S.)
* Correspondence: lidia.decastro@urjc.es

Abstract: In this paper, we analyse the impact of gender equality in managerial positions on wages and the gender wage gap in 22 European countries. We draw on the employer–employee microdata from the European Structure of Earnings Survey (E-SES) for the year 2018, which allows us to include firm fixed effects in our econometric specifications, thus controlling for both observed and unobserved heterogeneity at the firm level. The analysis is carried out not only at the mean but also across the wage distribution through unconditional quantile regressions. The results on the impact of gender equality in management on wages are mixed. However, we find that gender equality has a predominantly positive effect in the upper part of the wage distribution, and a negative effect in the middle and lower parts. The results on the impact on the gender wage gap show that in many cases, a more gender-equal management reduces the gender wage gap. Furthermore, gender equality in management reduces the gender wage gap mainly in the middle and lower part of the wage distribution.

Keywords: gender equality; management positions; wages; gender wage gap; quantile regression

Citation: de Castro Romero, Lidia, Víctor Martín Barroso, and Rosa Santero-Sánchez. 2023. Does Gender Equality in Managerial Positions Improve the Gender Wage Gap? Comparative Evidence from Europe. *Economies* 11: 301. https://doi.org/10.3390/economies11120301

Academic Editor: Fabio Clementi

Received: 8 November 2023
Revised: 8 December 2023
Accepted: 11 December 2023
Published: 14 December 2023

1. Introduction

In recent decades, gender equality has become a transversal international objective in all political, economic, labour and social aspects that govern the daily activities of countries (United Nations, SDG 5 on gender equality). Despite the progress made in terms of women's participation in the labour market, there are large disparities between countries, and important gaps remain. Both the glass ceiling and the gender pay gap persist as global issues, despite what Goldin (2014) describes as 'a great gender convergence' in human capital in much of the developed world.

Gender diversity in company management is considered a key objective in economic literature due to its potential impact on companies' economic and financial results, as well as its potential to initiate dynamics that reduce gender gaps in the workplace, in the conditions of access to positions of responsibility, in leaves of absence or in salaries. There is a large body of theoretical and empirical literature linking gender diversity in management to pay gaps, but they are without clear results. Despite there being no consensus on the key factors that facilitate the process to reduce the GWG and the magnitude of the effect, some authors such as Theodoropoulos et al. (2022) and Santero-Sánchez and Castro Núñez (2022) have pointed out that female managers reduce the GWG.

Gender diversity in decision making has traditionally been measured by the proportion of women in managerial positions. More than 6.7 million persons held a managerial position in the European Union out of the 27 Member States (EU): 4.3 million men (63% of all managers) and 2.5 million women (37%). At the EU level, this share increased slightly compared to 2012 (36%). Women account for less than a third of managers in Cyprus (19%), followed by Denmark (27%), Italy (28%), the Netherlands (29%), Czechia and Germany (both 31%) and Croatia (32%). Above average, Latvia is the only Member State where

women are the majority (53%) in this occupation. This is followed by Bulgaria (49%), Poland (48%), Slovenia (44%), Hungary and Sweden (both 42%) and Portugal (40%) (Eurostat 2020).

As far as the gender wage disparity is concerned, it persists throughout Europe. According to Eurostat (2023), the average GWG for the European Union (EU27) has narrowed slightly in recent years, falling from 14.5% in 2015 to 14.4% in 2018 and 12.7% in 2021. Even after controlling for observed factors that could explain the differences, such as skills and experience, the unexplained GWG (which is the literature's measure of wage discrimination) in the EU in 2018 showed a small variation, from an unadjusted GWG of 11.4% to an unexplained GWG of 11.4% (Leythienne and Pérez-Julián 2021). The countries with the largest unexplained gender gap in 2018 were Czechia (17.6%), Latvia (17.5%), Bulgaria (16.2%) and Croatia (15.2%), while the countries with the smallest gap were Sweden (7.6%), Cyprus (8.6%), Norway (8.7%) and Denmark (9.4%) (Leythienne and Pérez-Julián 2021).

This highlights the need for further analysis of this economic and social problem to clarify the mechanisms that can lead to more equal participation and opportunities. In this respect, the European Gender Equality Strategy 2020–2025 includes the promotion of gender equality and women's empowerment in the economy as one of its key areas (European Commission 2021). The aim of this initiative is to better understand the dynamics that promote and penalise gender equality in order to find tools within the reach of politics and organisational practices to achieve effective equality between men and women.

However, most analyses of the wage gap focus on quantifying it in average terms, which by default neglects the potential differences between workers at different points in the wage distribution, both at the lowest level, associated with the 'sticky floor' phenomenon, and at the top, associated with the 'glass ceiling'. Advances in wage gap decomposition techniques have made it possible to assess the wage gap at different points in the distribution. According to the International Labour Organization (International Labour Organization 2019), 44% of all 93 countries analysed have a higher GWG in management positions. Correspondingly, some research for some countries finds a higher GWG at the top of the pay scale, while for others, the results show higher pay gaps at the bottom of the pay scale.

The absence of agreement about the impact of gender diversity in management on the GWG points to the need to study this process in more detail. This effort would lead to a better understanding of whether different organisational practices have an impact or an equal impact at different pay percentiles, in the interest of designing and implementing more effective policies against the GWG.

Against this background, the objective of this study is to provide a more complete picture of the impact of gender diversity in decision-making positions on the GWG across the European pay distribution. Our approach uses a measure of management diversity that takes into account the impact of men's entry into companies with women in management positions, a perspective which is not sufficiently taken into account in measures based purely on the proportion of women in an organisation. The use of a broad group of European countries, in which heterogeneity is high, allows for the identification of country-specific conditions that help to understand the enablers and barriers that need to be considered when designing government and company policies aimed at reducing and eliminating the GWG.

Our analysis has been based on the microdata from the European Structure of Earnings Survey for the year 2018 provided by Eurostat. This database provides employer–employee matched data, which allows for the inclusion of establishment fixed effects in the econometric specifications, thus allowing us to account for both observed and unobserved establishment heterogeneity.

In order to analyse the impact of gender equality in managerial positions on wages and on the GWG, we have estimated an extended version of the traditional workhorse Mincerian wage equation. In addition, since gender equality in managerial positions may have different effects on wages and the GWG throughout the wage distribution, we have run unconditional quantile regressions following the methodology proposed by Firpo et al. (2009).

The remainder of the paper is organised as follows: Section 2 reviews the previous empirical literature. Section 3 describes our index of managerial gender equality and the econometric methodology used throughout the paper to estimate the impact of gender equality in management on wages and the GWG. Section 4 provides a description of the employer–employee dataset used in the empirical analysis and a description of the sample. Section 5 presents the results of the regression analyses, which are discussed in Section 6. The final section presents the conclusions.

2. Literature Review

Women's opportunities to hold leadership positions in organisations are restricted due to the glass ceiling, which contributes to the perpetuation of gender inequalities within organisations (Huffman 2016). According to social closure theory (Tomaskovic-Devey 1993), males monopolise advantages and limit women's opportunities for promotion, leading to a process of subordination.

Several hypotheses have been proposed to explain the under-representation of women in leadership roles in business and politics. These hypotheses range from demand-side constraints, including pre-existing social norms and gender stereotypes that create a 'glass ceiling' effect for women, to supply-side explanations, such as a reduction in women's working time. For instance, social norms and gender stereotypes may serve to bias managers and voters against recruiting women as managers and leaders (Huddy and Terkildsen 1993; Eagly and Karau 2002). Furthermore, insufficient interactions with women leaders may perpetuate biased perceptions of women's effectiveness in leadership (Beaman et al. 2009). Women themselves may not believe in their ability to lead because they rarely see other women in such positions (Beaman et al. 2012), and they may also leave high-level careers to have children (Bertrand et al. 2010).

A comparable analysis is found in the literature on GWG. Its existence has traditionally been explained by the decisions and motivations that differentiate women and men in their personal and professional careers. A supply-side approach emphasises the workers and gives them the responsibility for wage differentials. Becker's (1962) human capital theory provided the most important explanation of how choices about education, career interruptions, education or part-time work affect workers' productivity. Thus, the first explanations given for the GWG were those related to how women's lower human capital affected their wages (Mincer and Polachek 1977). Demand-side theories have also been developed and are more specific to the dynamics of the reproduction of discrimination by firms. The different roles assigned to women and men determine the most suitable jobs (horizontal segregation), the positions reserved for each (vertical segregation) and the working conditions, such as the value of each contribution in terms of wages (Akerlof and Kranton 2000; Conde-Ruiz and Marra de Artíñano 2016; Fortin 2005; Rubery et al. 2005).

However, the GWG is not homogeneous across the wage distribution. According to empirical literature (Hara 2018; Javdani 2015; Santero-Sánchez and Castro Núñez 2022), the GWG exhibits differences between the lower and upper ends of the wage distribution, which gives significance to the phenomena of 'sticky floors' and 'glass ceilings', respectively. Some research for some countries has found a higher GWG at the top of the pay scale (Cotter et al. 2001; Albrecht et al. 2003; Huffman 2004; Christofides et al. 2013; International Labour Organization 2019), while for others, the results have showed higher pay gaps at the bottom of the pay scale (Arulampalam et al. 2007; Christofides et al. 2013). Hence, a more profound comprehension of GWG dynamics requires a deeper analysis of wage differentials across the distribution (Huffman et al. 2017).

Establishing a connection between these two phenomena reveals the impact of women's under-representation in leadership positions on the GWG. According to the statistical discrimination theory (Phelps 1972), traditionally, the presence of men in decision-making positions and gender stereotypes have biased men's statistical information on the group to which they belong in order to infer women's productivity. Similarly, the discriminatory taste theory (Becker 1957), in line with social identity theory (Tajfel and Turner 1979),

explained that men value women less because they belong to different social groups. As a consequence, men hold biases against women that impact their pay; they employ or promote women when their lower pay neutralises the disutility they experience. Therefore, the consequences of this discrimination are more significant as women move closer to the top of the hierarchy, which is associated with higher salaries.

Thus, the historical connection between formalised pay systems and gender pay inequality can be attributed, in part, to the disproportionate representation of men in management roles (Abraham 2017), whereas a higher GWG found in the upper percentiles can frequently be linked to salary supplements received, which reward greater effort in terms of time, travel and the volume of projects undertaken (De la Rica et al. 2015; Christofides et al. 2013). Therefore, many pay equity programmes focus on formalising the salary assignment to avoid discriminatory situations due to the discretion of decision-makers. However, other research has found that a higher proportion of female managers with discretionary pay-setting power is negatively related to the GWG (Theodoropoulos et al. 2022; Abraham 2017).

Based on the above, some empirical studies have linked gender diversity in management to the promotion of equal conditions in a firm's work environment. Several have suggested that diversity in management reduces barriers to career advancement (Dalvit et al. 2021; Kunze and Miller 2017; Mateos de Cabo et al. 2011; Matsa and Miller 2011). And others have referred to equality as a method to reduce the GWG (Bell 2005; Bertrand et al. 2019; Cardoso and Winter-Ebmer 2010; Cohen and Huffman 2007; Flabbi et al. 2019; Theodoropoulos et al. 2022; Hensvik 2014; Hirsch 2013; Vega et al. 2016). However, the results were quite varied in terms of significance and the size of the specific group of women affected. Furthermore, the impact on the GWG is not homogeneous across the distribution. Huffman et al. (2017) conducted studies in Germany, and Santero-Sánchez and Castro Núñez (2022) in Spain, and they concur that policies on promoting women decrease wage inequality, particularly in the lower ranks. Meanwhile, Bertrand et al. (2019) found that increasing women's representation on boards was associated with a significant decrease in the GWG among top executives but found limited evidence of changes at the bottom of the corporate hierarchy.

Two mechanisms can potentially explain the impact of female managers on the GWG within an organisation. Firstly, female staff may benefit from the homophily and mentoring provided by female managers when interacting with them (Hultin and Szulkin 2003). Secondly, female managers might use their organisational power to change organisation practices, resulting in a more gender equitable organisation (Cohen and Huffman 2007). The results of Zimmermann (2022) showed a decrease in the GWG only with female executives at the second level. They are able to redistribute wages between men and women because they are in direct contact with workers. Meanwhile, first-level managers, who could change the organizational structure by making strategic decisions, only affected GWG when firms were small enough to contact workers.

An alternative possibility is that the presence of women in management does not have any impact on the GWG because the measures adopted take on a purely symbolic meaning and do not permeate the organisational culture (Huffman 2016). Moreover, several studies, including the ones conducted by Magda and Cukrowska-Torzewska (2019) and by van Hek and van der Lippe (2019) for a set of European Union countries, found no relationship between the proportion of female managers and the income of women and men. In this context, female employees would essentially be cogs in the machine (Cohen and Huffman 2007) without changing the discriminatory dynamics of organisations.

Furthermore, if women have not achieved sufficient power to change organisational practices, their presence will have no impact on the GWG. However, some women leaders may not be changing agents simply because they do not want to be. Scholars have suggested that it is women who accept the status quo who are promoted to positions of responsibility (Cohen and Huffman 2007; Huffman 2016), and that they sometimes follow this behaviour because they assume gender stereotypes to be true (Derks et al. 2011). In

contrast, according to the queen bee theory, other female managers would not oppose gender discrimination because they believe it would risk their position. As a result, women managers may benefit outgroup members (men) by imitating their behaviours to fit in (Kalogeraki and Georgakakis 2021) in an attempt to justify and protect their position in a male-dominated environment.

Lastly, one factor to consider is the market's sectoral segregation, which allows women to be promoted to positions of responsibility in companies with a majority of women and belonging to a feminised sector. It is acknowledged that sectors with a high proportion of women are devalued in terms of salaries. This phenomenon could be explained through the theory of structural occupational crowding (Groshen 1991; Sorensen 1989) and the theory of the devaluation of women's work (England 1992). According to the former, women have a propensity for lower-paying positions and therefore are crowded out of higher-paying jobs. The latter argues that female-dominated occupations have a lower wage value and/or occupational prestige than male-dominated occupations simply because they are occupied by women.

In line with the previous point, the results obtained by Santero-Sánchez and Castro Núñez (2022) for Spain show that increasing gender diversity in managerial positions only had a significant impact on reducing the GWG when more women access these positions in companies managed by men. Thus, it is crucial to examine the association between GWG and diversity in company management from an analytical view.

In conclusion, the existing literature outlines the causes of both the glass ceiling and the GWG and links them; the GWG refers to another barrier under the term glass ceiling (Bertrand et al. 2019), whereas the under-representation of women in decision-making positions leads men to decide the wages of subordinate women with a gender bias. As such, the presence of women in leadership positions is theoretically associated with a GWG reduction, as they can form networks to support other women and influence gendered-bias organisational structures. Thus, an analysis of the impact of the presence of women in decision making on the GWG is necessary. And it is of analytical interest to understand its effect not only on the median wage but also on the various levels of the wage distribution because the GWG varies depending on wage levels.

3. Methodology

3.1. Measuring Gender Equality in Managerial Positions

To measure the degree of gender equality in managerial positions, we calculate the following index[1] for each establishment in the sample:

$$GEMP_j = 1 - \left| \frac{X_j^F}{X_j^T} - \frac{X_j^M}{X_j^T} \right| \quad (1)$$

where X_j denotes the number of managers in an establishment j, and the superscripts F, M and T relates to female, male and total. Thus, the gender equality index in managerial positions (*GEMP*) is calculated for each workplace, as one minus the difference in absolute value in the share of female and male managers. Our index is gender neutral since all gaps, regardless of whether they are to the disadvantage of women or men, are taken into consideration and treated in the same way. The index takes values in the interval [0,1], where 1 represents gender equality (equal share of female and male managers) and 0 represents inequality.

3.2. Baseline Estimates

Following common practice on previous empirical works analysing the gender wage gap, our departure point is an augmented version of the traditional workhorse Mincerian wage equation (Mincer 1974). This equation rests on the premise that the wage of a worker is determined by their productivity, which, in turn, depends on the level of education

and the accumulated work experience of the worker. Our augmented version takes the following form:

$$W_{ij} = \alpha + \beta female_{ij} + \delta GEMP_j + \lambda female_{ij} \times GEMP_j + \sum_k \gamma_k Z_{kij} + \varepsilon_{ij} \quad (2)$$

where the subscript i refers to workers and the subscript j refers to local units or establishments. The dependent variable (W_{ij}) is the hourly wage log-transformed. Among the explanatory variables, *female* is a dummy variable which takes a value of one if worker i is female and zero otherwise, *GEMP* is the value of the gender equality index in management positions in establishment j, and Z is a vector of explanatory variables including worker (sex, age and education level), job (tenure, type of contract, type of working day and occupation) and establishment characteristics (side, location and economic sector). Lastly, ε_{ij} is the idiosyncratic error term.

As for the coefficients of interest to be estimated, the value of β represents the wage difference in log points between females and males with similar individual and job characteristics working in similar establishments. The adjusted gender wage gap is calculated from the estimated β coefficient as GWG = $(e^{\beta} - 1) \times 100$. The coefficient δ quantifies the effect of gender equality in management on wages. More specifically, its value represents the impact on log wages of a one-unit increase in the gender equality index in management. Lastly, the parameter λ of the interaction term between gender equality in management and the female dummy measures the effect of gender equality in management on the adjusted GWG.

Additionally, we include establishment fixed effects which allows us to control for unobserved heterogeneity between local units or firms. The equation to be estimated now is

$$W_{ij} = \alpha + \beta female_{ij} + \lambda female_{ij} \times GEMP_j + \sum_k \gamma_k Z_{kij} + \mu_j + \varepsilon_{ij} \quad (3)$$

where μ_j is the establishment fixed effect. In Equation (3), we include neither the gender equality index in management nor the characteristics of the establishment since they are absorbed by the workplace fixed effect. While the estimation results are improved, a drawback of using fixed effects is that we cannot address the effect of gender equality on wages.

3.3. Unconditional Quantile Regression

To analyse the impact of gender equality in management positions on wages and on the GWG throughout the wage distribution, we conducted UQR following the methodology proposed by Firpo et al. (2009). In contrast to conditional quantile regression, UQR defines quantiles for the variable of interest (in our case, the log hourly wage) before conducting regression. This implies that when covariates are included in the regression, they serve to account for their impact on the specific relationship under study (such as the relationship between wages, gender and equality in management positions). However, the inclusion of covariates does not influence the assignment of observations to specific quantiles within the wage distribution (Killewald and Bearak 2014). UQR is a two-step procedure where, in the first step, the log hourly wage variable is transformed into the recentred influence function (RIF) of the unconditional quantile of the wage distribution, defined as

$$RIF(W_i; Q_\tau) = Q_\tau + \frac{\tau - 1[W_i \leq Q_\tau]}{f_W(Q_\tau)} \quad (4)$$

where $\tau \in (0, 1)$ is a given quantile, Q_τ is the value of the wage variable (W_i) at the τth quantile, $f_W(Q_\tau)$ is the density function of wages at quantile Q_τ and $1[W_i \leq Q_\tau]$ is a dummy variable indicating whether the wage observation is at or below quantile Q_τ. In the second step, standard OLS regression can be applied to Mincerian wage equations, substituting the dependent variable (log hourly wage) with the RIF calculated at different quantiles. Thus,

for a given quantile $\tau \in (0, 1)$, the UQR equivalent to Equations (2) and (3) are represented, respectively, by the following equations[2]:

$$\hat{RIF}(W_i; Q_\tau) = \alpha + \beta female_{ij} + \delta GEMP_j + \lambda female_{ij} \times GEMP_j + \sum_k \gamma_k Z_{kij} + \varepsilon_{ij} \quad (5)$$

$$\hat{RIF}(W_i; Q_\tau) = \alpha + \beta female_{ij} + \lambda female_{ij} \times GEMP_j + \sum_k \gamma_k Z_{kij} + \mu_j + \varepsilon_{ij} \quad (6)$$

4. Data and Sample Description

This research paper draws upon the employer–employee microdata from the European Structure of Earnings Survey (E-SES) for the year 2018. The survey, which is a comprehensive source of labour market information, is conducted on a quadrennial basis through a collaborative effort between Eurostat and the National Statistical Institutes of numerous European nations, ensuring a harmonised and standardised methodology across the participating countries. Its design employs a two-stage random sampling procedure. Initially, a stratified random sample of firms and/or local units is selected, and then, a random sample of employees is drawn from these chosen units. To minimise potential biases in the data, the survey is typically conducted in the month of October, which is usually associated with fewer job absences due to annual leave or public holidays. The E-SES encompasses enterprises and local units, both public and private, with a workforce of at least 10 employees, spanning economic activities from section B to S (excluding O) as defined in NACE Rev. 2. Hence, the survey does not include information on workers from section A (agriculture, forestry and fishing), section T (activities of households as employers) and section U (activities of extraterritorial organisations and bodies).

Since the survey provides employer–employee matched data, individuals working in the same establishment can be identified, a fact that enables us to include local-unit or establishment fixed effects in the econometric specifications. Thus, it is possible to account for both establishment observed and unobserved heterogeneity, thus improving the econometric estimates. Nonetheless, the survey presents two noteworthy limitations. On the one hand, information in certain key factors to explain wage disparities among individuals, such as marital status and the number of children, is missing. Secondly, the E-SES lacks a longitudinal design, which means that it does not capture the trajectory of employers and employees over time.

The key variable in our analysis is the hourly wage, which is directly available within the E-SES dataset. We harmonised the variable to euros for nations that do not use the euro and whose wage variable was originally denominated in their national currencies, enhancing the interpretability and comparability of wage data across all countries. Average exchange rates for the year 2018 published by the European Central Bank were used in the transformation.

Workers in management positions are identified as those with occupation defined as 'Group A' in the National Classification of Occupations for 2011, which includes directors and managers.

As indicated in the methodology section, we use three sets of covariates as explanatory variables in our wage equations: individual or worker characteristics, job characteristics, and establishment characteristics. Regarding the set of individual characteristics, we include information on age as a proxy of work experience and the level of education. Data on age were aggregated into six categories, as follows: workers below the age of 19, between 20 and 29, between 30 and 39, between 40 and 49, between 50 and 59, and 60 years old and above. For the level of education, the original four categories in the E-SES were kept, as follows: basic education, secondary education, tertiary education (up to 4 years) and tertiary education (more than 4 years). As for the set of job characteristics, we included information on job tenure in years (log transformed), the type of contract (permanent or fixed-term), the type of working day (full-time or part-time), and occupation aggregated at the nine major groups according to ISCO-08[3]. Regarding the characteristics of the establishment or local unit, we utilised the accessible data regarding size, specifically

the number of workers, geographical location at the NUTS 1 level, and the economic sector. Regarding establishment size, we relied on enterprise size information, as specific data for individual local units is often optional and consequently unavailable for numerous countries. We grouped establishments into three categories based on size, as follows: those with fewer than 50 workers, those with between 50 and 249 workers, and those with 250 workers or more. Regarding the economic sector, we included the section as defined in NACE Rev. 2.

In all estimates, we use the worker-level weights provided with the survey to account for sample selection probabilities, as well as to address any discernible non-response bias. This approach guarantees the representativeness of our findings for all individuals within the surveyed population.

We restricted the sample of local units to those that provided data relating to at least two workers holding a managerial role. As a result, we excluded establishments with only one managerial position due to the absence of gender equality. That is, those establishments cannot demonstrate gender diversity in management positions. Furthermore, the selection of local units with a minimum of two workers ensures the accurate identification of establishment fixed effects in the econometric estimates (Casado-Díaz et al. 2020). The drawback of this decision is that many small firms were not included in the analysis, resulting in a certain sample selection bias.

Given the availability of data, our ultimate dataset comprises comprehensive details regarding firm and worker attributes across 22 countries, all of which are members of the European Union except for Norway. Summary sample statistics are shown in Table 1. The sample size of workers varies considerably across countries, ranging from 3382 individuals in Greece to 2,153,729 in Czechia. The former, along with Norway (1,340,799) and Denmark (1,285,307), have the largest sample size. Among the countries where the mean index of gender equality in management has the highest values are Romania (65.4), Sweden (53.5), Croatia (52.6), France (51.4) and Malta (50.8). This evidence is consistent with the evolution of the EU Gender Equality Index (GEI) between 2010 and 2018 (Eurofound and EIGE 2021) and with the evolution of the score in the domain of power, which includes gender balance on company boards. Sweden leads in the overall GEI score, and Malta is one of the countries with the highest growth in the overall index (over 16% in the period). In both Sweden and Romania, women's share of women's employment accounted for by education, health and welfare activities is about four times that of men. And France is one of the countries that has increased its power domain as a result of quota policies.

Table 1. Sample statistics.

Country	Code	Nº Workers	Nº of Local Units	Gender Equality in Management Positions (Mean)	Mean Hourly Wage		GWG (%)
					Male	Female	
Belgium	BE	13,744	339	39.9 (40.8)	26.7 (10.1)	24.6 (10.0)	8.0
Bulgaria	BG	82,561	1761	48.9 (37.8)	4.4 (4.0)	3.6 (2.8)	18.7
Cyprus	CY	17,927	127	41.6 (34.6)	16.2 (14.0)	13.8 (9.3)	14.6
Czechia	CZ	2,153,729	7353	45.3 (32.4)	8.4 (5.6)	6.6 (3.4)	21.5
Germany	DE	120,910	3727	28.6 (38.8)	31.5 (18.5)	23.9 (11.1)	24.2
Denmark	DK	1,285,307	15,476	47.9 (35.8)	34.8 (17.2)	29.2 (11.4)	15.9
Estonia	EE	22,437	1046	34.2 (40.9)	8.5 (5.3)	6.8 (4.1)	20.4
Greece	EL	3382	189	42.4 (38.2)	16.1 (17.2)	11.5 (8.8)	28.1
Spain	ES	16,620	1063	46.1 (42.1)	18.7 (11.5)	16.2 (8.8)	13.2
France	FR	125,060	8790	51.4 (37.2)	23.8 (24.8)	19.2 (13.2)	19.0
Croatia	HR	21,370	621	52.6 (40.2)	7.3 (5.5)	6.6 (4.5)	9.3
Hungary	HU	606,293	7307	48.7 (36.8)	6.5 (5.2)	5.5 (3.1)	16.3
Italy	IT	12,805	479	40.1 (41.4)	21.9 (23.6)	17.7 (10.1)	19.1

Table 1. *Cont.*

Country	Code	N° Workers	N° of Local Units	Gender Equality in Management Positions (Mean)	Mean Hourly Wage		GWG (%)
					Male	Female	
Latvia	LV	139,945	3247	47.3 (37.3)	7.3 (6.0)	6.0 (4.5)	17.0
Malta	MT	31,161	227	50.9 (37.5)	14.1 (10.5)	12.2 (5.5)	13.2
Netherlands	NL	54,625	1243	46.7 (36.1)	22.6 (15.1)	18.7 (11.9)	17.4
Norway	NO	1,340,799	31,189	46.7 (36.1)	35.0 (71.5)	29.6 (61.4)	15.5
Poland	PL	646,405	11,546	49.3 (34.3)	7.1 (5.9)	6.1 (4.1)	14.3
Portugal	PT	14,337	548	48.0 (41.9)	11.8 (10.4)	9.5 (7.0)	19.5
Romania	RO	179,634	7266	65.4 (35.3)	6.1 (5.5)	6.0 (4.7)	1.6
Sweden	SE	229,130	2080	53.5 (29.3)	22.6 (11.5)	19.3 (7.0)	14.5
Slovakia	SK	811,412	4291	47.2 (33.9)	8.0 (7.0)	6.2 (3.6)	22.1

Notes: Grossing-up factors were used to compute gender equality in management, the mean hourly wages, and the raw gender wage gap (GWG). The hourly wage was expressed in euros for countries that do not use the euro and whose the wage variable was originally measured in their national currency by using the average exchange rates in 2018 published by the European Central Bank. The GWG is calculated as the difference between male and female mean hourly wages divided by the male mean hourly wage. Source: own elaboration based on the authors' estimate using E-SES 2018.

As it can be noted, females are underpaid compared to men in all 22 countries in the sample. The gender pay gap varies from 1.6% (Romania) to 28.1% (Greece), yet it persists at a rate higher than 10% in all countries except for three. Romania is an exception because jobs in the public sphere are, on average, financially more rewarding that in the private sector, thus causing a reduction in the GWG (Leythienne and Pérez-Julián 2021). Apart from Romania, the countries with the smallest gender pay gap are Belgium (8%), Croatia (9.3%) and Spain (13.2%). On the other hand, following Greece, the countries with the largest gender gap are Germany (24.2%), Slovakia (22.1%) and Czechia (21.2%).

Additionally, Figure 1 plots each country in the sample, with the unadjusted gender pay gap on the y-axis and the average of the GEMP index on the x-axis. As can be seen, the cloud of points shows a certain negative trend, indicating that those countries where the GEMP index is higher on average have a lower GWG.

Figure 1. Gender equality in management and the raw GWG. Source: own elaboration based on the authors' estimate using E-SES 2018.

5. Econometric Results

5.1. Baseline Estimates

The results of Equation (2) are shown in Table 2. For each country, the estimated coefficients and their respective standard error deviations for the variables female, *GEMP* and the cross product of both are displayed[4]. Additionally, the number of observations and the resulting adjusted R^2 are presented.

Table 2. Summary of wage equation estimates by OLS.

Country	Female		Gender Equality in Management (GEM)		Female × GEM		N	R^2
	Coefficient	(s.e.)	Coefficient	(s.e.)	Coefficient	(s.e.)		
BE	−0.0136	(0.0083)	0.0146 *	(0.0081)	0.0191	(0.0172)	13,744	0.86
BG	−0.1636 ***	(0.0061)	0.0158 **	(0.0070)	0.0702 ***	(0.0094)	82,561	0.43
CY	−0.1621 ***	(0.0125)	−0.0060	(0.0228)	0.0996 ***	(0.0261)	17,927	0.74
CZ	−0.1721 ***	(0.0022)	−0.0336 ***	(0.0034)	−0.0055	(0.0038)	2,153,729	0.52
DE	−0.0861 ***	(0.0039)	−0.0063	(0.0056)	−0.0088	(0.0072)	120,910	0.60
DK	−0.1068 ***	(0.0011)	0.0053 ***	(0.0013)	0.0094 ***	(0.0017)	1,285,307	0.56
EE	−0.1901 ***	(0.0143)	−0.0976 ***	(0.0202)	0.0499 *	(0.0257)	22,437	0.34
EL	−0.0964 ***	(0.0294)	−0.0021	(0.0354)	−0.0449	(0.0470)	3382	0.53
ES	−0.1335 ***	(0.0159)	0.0092	(0.0147)	0.0103	(0.0217)	16,620	0.49
FR	−0.1311 ***	(0.0077)	0.0173 **	(0.0072)	−0.0150	(0.0108)	125,060	0.46
HR	−0.1620 ***	(0.0093)	−0.0165 *	(0.0092)	0.0006	(0.0135)	21,370	0.48
HU	−0.0958 ***	(0.0036)	−0.0031	(0.0048)	0.0098 *	(0.0055)	606,293	0.58
IT	−0.1496 ***	(0.0113)	−0.0503 ***	(0.0151)	0.0834 ***	(0.0178)	12,805	0.50
LV	−0.2117 ***	(0.0090)	−0.0280 **	(0.0119)	0.0855 ***	(0.0145)	139,945	0.39
MT	−0.0883 ***	(0.0189)	−0.0463 **	(0.0234)	−0.0205	(0.0286)	31,161	0.48
NL	−0.0640 ***	(0.0061)	0.0301 ***	(0.0078)	−0.0126	(0.0091)	54,625	0.71
NO	−0.1029 ***	(0.0009)	−0.0058 ***	(0.0010)	0.0102 ***	(0.0013)	1,340,799	0.51
PL	−0.1654 ***	(0.0019)	−0.0068 ***	(0.0023)	0.0127 ***	(0.0030)	646,405	0.48
PT	−0.1497 ***	(0.0131)	−0.0401 **	(0.0158)	0.0190	(0.0213)	14,337	0.70
RO	−0.0897 ***	(0.0057)	0.0098 *	(0.0053)	−0.0465 ***	(0.0074)	179,634	0.53
SE	−0.0565 ***	(0.0032)	0.0140 ***	(0.0039)	−0.0092 *	(0.0048)	229,130	0.54
SK	−0.2114 ***	(0.0049)	−0.0357 ***	(0.0073)	0.0282 ***	(0.0087)	811,412	0.44

Notes: Sampling weights are used in all estimations. Robust standard errors (s.e.) between parentheses. All estimations include a constant term. Statistically significance indicated by *** 1% level, ** 5% level and * 10% level. Source: own elaboration based on the authors' estimate using E-SES 2018.

Overall, the results presented provide evidence for the existence of a relatively high GWG in almost all the countries considered. Furthermore, we find that the *GEMP* has a positive effect on wages in 10 countries, while a higher *GEMP* is also associated with a lower GWG in 10 out of 22 countries.

The estimated coefficient for the female variable is negative and statistically significant at the 1% level in all countries except for Belgium. Its value ranges from −0.0565 log points (Sweden) to −0.2117 log points (Latvia), which implies an adjusted gender wage gap ranging from 5.5% to 19.9%.

The relationship between gender equality in management positions and wages yields mixed results. In 10 of the 22 countries examined, a higher gender equality in management corresponds to lower wages, as indicated by the negative and statistically significant coefficient for the *GEMP* variable at conventional levels. Conversely, in 6 of the 22 countries, the estimated coefficient for the *GEMP* variable is positive and statistically significant. Thus, in Belgium, Bulgaria, Denmark, France, the Netherlands and Romania, there is evidence that a greater level of gender equality in management is associated with increased wages. However, for the other six countries, the estimated coefficient of *GEMP* is not statistically significant. Empirical studies have showed that firms with higher gender diversity in management positions improve compensation conditions for all workers (Santero-Sánchez and Castro Núñez 2022 for Spain).

Regarding the relationship between gender equality in managerial positions and the GWG, the results are also mixed. In 10 out of the 22 countries, the estimated coefficient for the cross product of the variables female and *GEMP* is positive and statistically significant, indicating that an increase in gender equality in management positions leads to a reduction in the GWG, as supported by most of the previous empirical literature (Bell 2005; Cardoso and Winter-Ebmer 2010; Hensvik 2014). These countries are Bulgaria, Cyprus, Denmark, Estonia, Hungary, Italy, Latvia, Norway, Poland and Slovakia. Among them, those where the estimated coefficient have greater values are Cyprus (0.0996 log points), Latvia (0.0855 log points), Italy (0.0834 log points) and Bulgaria (0.0702 log points). These values imply that a one-point increase in the index of *GEMP* (e.g., the difference between a female- or male-led establishment and an establishment where management is gender neutral) is associated with a decrease in the GWG of almost 9 percentage points in Cyprus (from 15% to 6.1%), 7.2 percentage points in Latvia (from 19.1% to 11.9%), 7.5 percentage points in Italy (from 13.9% to 6.4%) and in 6.2 percentage points in Bulgaria (from 15.1% to 8.9%). In the other six countries where the coefficient for the cross product of the variables female and *GEMP* is positive and statistically significant, the impact is smaller, ranging from a reduction in GWG of 0.8 percentage points in Denmark (from 10.1% to 9.3%) to a reduction of 4.2 percentage points in Estonia (from 17.3% to 13.1%). For the remaining countries, in 10 cases, the coefficient is not statistically significant at conventional levels, indicating that gender equality in management positions has no effect on the GWG. Spain is one of these countries, and this result coincides with Santero-Sánchez and Castro Núñez (2022). In only two cases, the estimated coefficient is negative and statistically significant: Romania and Sweden.

The estimation results of Equation (3), where the wage equation is estimated including establishment or local unit fixed effects, are shown in Table 3. Again, we find a relatively high GWG in all the countries in the sample. Also, in eight countries, the results suggests that a higher GEMP reduces the GWG.

Table 3. Summary of wages equation estimates by OLS with establishment fixed effects.

Country	Female		Female × GEM		N	R^2
	Coefficient	(s.e.)	Coefficient	(s.e.)		
BE	−0.0136 *	(0.0081)	0.0180	(0.0190)	13,744	0.87
BG	−0.1104 ***	(0.0049)	0.0309 ***	(0.0077)	82,561	0.74
CY	−0.1157 ***	(0.0134)	0.0461 *	(0.0249)	17,927	0.79
CZ	−0.1479 ***	(0.0020)	0.0063 *	(0.0034)	2,153,729	0.69
DE	−0.1195 ***	(0.0048)	0.0230 ***	(0.0079)	120,910	0.73
DK	−0.0890 ***	(0.0010)	0.0020	(0.0016)	1,285,307	0.66
EE	−0.1109 ***	(0.0132)	0.0259	(0.0217)	22,437	0.65
EL	−0.0767 ***	(0.0246)	−0.0431	(0.0391)	3382	0.70
ES	−0.1222 ***	(0.0136)	0.0227	(0.0187)	16,620	0.69
FR	−0.1171 ***	(0.0070)	−0.0234 **	(0.0101)	125,060	0.61
HR	−0.0991 ***	(0.0089)	−0.0346 **	(0.0137)	21,370	0.69
HU	−0.0897 ***	(0.0033)	0.0177 ***	(0.0049)	606,293	0.75
IT	−0.1196 ***	(0.0113)	0.0366 **	(0.0174)	12,805	0.62
LV	−0.0878 ***	(0.0083)	−0.0380 ***	(0.0131)	139,945	0.66
MT	−0.1122 ***	(0.0192)	0.0317	(0.0278)	31,161	0.59
NL	−0.0657 ***	(0.0059)	−0.0004	(0.0092)	54,625	0.76
NO	−0.0728 ***	(0.0009)	−0.0137 ***	(0.0014)	1,340,799	0.64
PL	−0.1167 ***	(0.0018)	−0.0121 ***	(0.0030)	646,405	0.69
PT	−0.1447 ***	(0.0123)	0.0466 **	(0.0202)	14,337	0.79
RO	−0.0950 ***	(0.0058)	−0.0266 ***	(0.0073)	179,634	0.72
SE	−0.0647 ***	(0.0033)	0.0092 *	(0.0051)	229,130	0.62
SK	−0.1336 ***	(0.0042)	−0.0102	(0.0070)	811,412	0.67

Notes: Sampling weights are used in all estimations. Robust standard errors (s.e.) between parentheses. All estimations include a constant term. Statistically significance indicated by *** 1% level, ** 5% level and * 10% level. Source: own elaboration based on the authors' estimate using E-SES 2018.

Here, the estimated coefficient for the female variable provides an estimate of the GWG within establishments or local units. The estimated coefficient is negative and statistically significant at the 1% level in all 22 countries. Compared to the results of Equation (2), the GWG is lower in 16 countries. Its value now ranges from −0.0136 log points (Belgium) to −0.1479 log points (Czechia), implying an adjusted gender wage gap ranging from 1.4% to 13.7%. The countries where the GWG decreased more when including establishment fixed effects are Latvia (from 19.1% to 8.4%), Estonia (from 17.3% to 10.5%) and Slovakia (from 19.1% to 12.5%).

The estimated coefficient for the cross product of the female and *GEMP* variables is now positive and statistically significant in eight countries (Bulgaria, Cyprus, Czechia, Germany, Hungary, Italy, Portugal and Sweden), suggesting that an increase in gender equality in managerial positions leads to a reduction in the GWG. Previous studies by Hirsch (2013) for Germany, Cardoso and Winter-Ebmer (2010) for Portugal and Hensvik (2014) for Sweden have reached the same result. The countries where the estimated coefficient have greater values are Portugal (0.0466 log points), Cyprus (0.0461 log points), Italy (0.0309 log points) and Bulgaria (0.0309 log points). A further interpretation of these coefficients suggests that a one-point increase in *GEMP* within the same establishment or local unit would reduce the GWG in 4.1 percentage points in Portugal (from 13.5% to 9.3%), 4.2 percentage points in Cyprus (from 10.9% to 6.7%), 3.3 percentage points in Italy (from 11.3% to 8%) and 2.8 percentage points in Bulgaria (from 10.5% to 7.6%). As for the rest of the countries, the estimated coefficient for the cross product of the female and *GEMP* variables is not statistically significant in 11 countries (including Spain, with the same result in Santero-Sánchez and Castro Núñez (2022)) and is positive and significant at conventional levels in three countries: France, Hungary, and Latvia.

5.2. Unconditional Quantile Regression Results

Table 4 summarises the results of the UQR without including establishment fixed effects, as stated in Equation (5). For each country, the table reports the estimated coefficients and their respective standard deviations for the variables female, *GEMP* and the cross product of both, in the 10th, 25th, 50th, 75th, and 90th percentiles of the wage distribution.

In summary, three main results are obtained: (i) the GWG is lower at lower percentiles and higher at the top of the wage distribution, (ii) GEMP has a positive effect on the upper part of the wage distribution and a negative effect at the middle and lower parts of the distribution, and (iii) GEMP reduces the GWG in the middle and lower part of the wage distribution.

The results indicate that the GWG is lower in the lower part of the wage distribution and then rises progressively as we move up the wage distribution. This fact is clearly observed in the majority of countries, with the exception of Belgium, Cyprus, Malta and Spain. The countries where the difference in the GWG at the upper and lower part of the wage distribution is larger are Hungary, with a difference of almost 21 percentage points (21.1% at the 90th percentile and 0.5% at the 10th percentile); Estonia, with a difference of 18.7 percentage points (24.1% at the 90th percentile and 5.4% at the 10th percentile); and Latvia, with a difference of 15.4 percentage points (26.1% at the 90th percentile and 10.6% at the 10th percentile). This finding provides evidence for the well-known fact that women experience a glass ceiling in the job market (Arulampalam et al. 2007; Chzhen and Mumford 2011; Christofides et al. 2013; Blau and Kahn 2017; Gharehgozli and Atal 2021; Santero-Sánchez and Castro Núñez 2022). As Bertrand et al. (2019) point out, there is an under-representation of women at the top of the labour market, and wage gender gaps are larger there than average, which is often referred to as the glass ceiling, meaning that women are under-represented (or over-represented) in high (or low) paying jobs, and this under-representation becomes more noticeable as we move up the wage distribution. As a result, the wage difference between male and female workers are greater at the top of the wage distribution than in the middle or at the bottom.

When examining the effect of *GEMP* on wages, the results show that gender equality in management has a predominantly positive effect in the upper part of the wage distribution, and a negative effect at the middle and lower parts of the distribution. Indeed, the estimated coefficient for *GEMP* is positive and statistically significant in 12 and 15 countries at the 75th and 90th percentile, respectively. On the other hand, the number of countries where the estimated coefficient is negative and statistically significant at the 10th, 25th and 50th percentiles is 15, 14, and 16, respectively. Furthermore, countries can be categorised into five groups. In the first group with the Netherlands and Sweden, a higher *GEMP* is associated with higher wages in almost all points of the wage distribution. Moreover, the effect is higher at the top than at the bottom. In the Netherlands, a one-point increase in the *GEMP* index is associated with an increase in wages of 1.4 percentage points at the 10th percentile, and an increase of 5.3 percentage points at the 90th percentile. In Sweden, a one-point increase in the index of *GEMP* is associated with an increase of 1.1 percentage points in wages at the 10th percentile and an increase of 2.7 percentage points at the 90th percentile. In the second group, consisting of Bulgaria, Cyprus, Denmark, France, Germany, Hungary, Norway, Portugal and Romania, the estimated coefficient of the *GEMP* variable is positive and statistically significant in the upper part of the wage distribution (75th and 90th percentiles) and negative and statistically significant in the middle and lower parts (50th, 25th and 10th percentiles). Thus, high-paid workers benefit from more gender-equal management, while middle-paid and lower-paid workers are negatively affected in these countries. In the third group, formed by Belgium, Poland and Slovakia, the estimated coefficient for the *GEMP* variable is positive and statistically significant only at the extremes of the distribution (10th and 90th percentiles) and negative and significant for the rest. On the other hand, in Spain (fourth group), the wage of highly paid workers (90th percentile) and low-paid workers (10th percentile) is negatively affected by a higher gender equality in management, while the wage of workers in the rest of the distribution is positively affected. Finally, in the fifth group, consisting of Croatia, Czechia, Estonia, Greece, Italy, Latvia and Malta, the estimated coefficient for *GEMP* is negative and statistically significant in almost the entire wage distribution.

With regard to the effect of *GEMP* on the GWG, the results suggest that gender equality in managerial positions reduces the GWG mainly in the middle and lower part of the wage distribution. More specifically, the estimated coefficient of the cross product of the female and *GEMP* variables is positive (negative) and statistically significant at conventional levels in 11 countries (6 countries) at the 10th percentile, 17 countries (2 countries) at the 25th percentile, and 17 countries (4 countries) at the 50th percentile. However, at the top of the wage distribution, a greater gender equality in management is associated with a higher GWG. The estimated coefficient for the cross product of the female and *GEMP* variables is negative (positive) and statistically significant in 10 countries (8 countries) at the 75th percentile and in 12 countries (4 countries) at the 90th percentile. The sample of countries can also be divided into five groups. The first group, in which the estimated coefficient for the cross product of the female and *GEMP* variables is now positive and statistically significant in almost the entire distribution, consists of Belgium, Bulgaria, Estonia, Italy, Latvia and Norway. Thus, in these six countries, the results indicate that higher gender equality in management reduces the GWG at different points of the wage distribution. In the case of Bulgaria, Italy and Latvia, the results also show that the impact is larger at the middle (50th percentile) and the lower (10th and 25th percentiles) parts of the distribution than at the top (75th and 90th percentiles).

For illustrative purposes, Figure 2 shows the reduction in the GWG (in percentage points) resulting from a one-point increase in the *GEMP* index. For example, in Italy, the difference between an establishment where managers are all male or all female and an establishment where management is gender neutral, implies a reduction in the GWG of 7.7 percentage points at the 10th percentile and of 11.4 percentage points at both the 25th and 50th percentiles, compensating for the overall underpayment suffered by women at these three percentiles. However, at the 75th percentile, the GWG is reduced by 3.1 percentage

points (from 14.6% to 11.6%) and at the 90th percentile by 2.4 percentage points (from 21.5% to 19.1%).

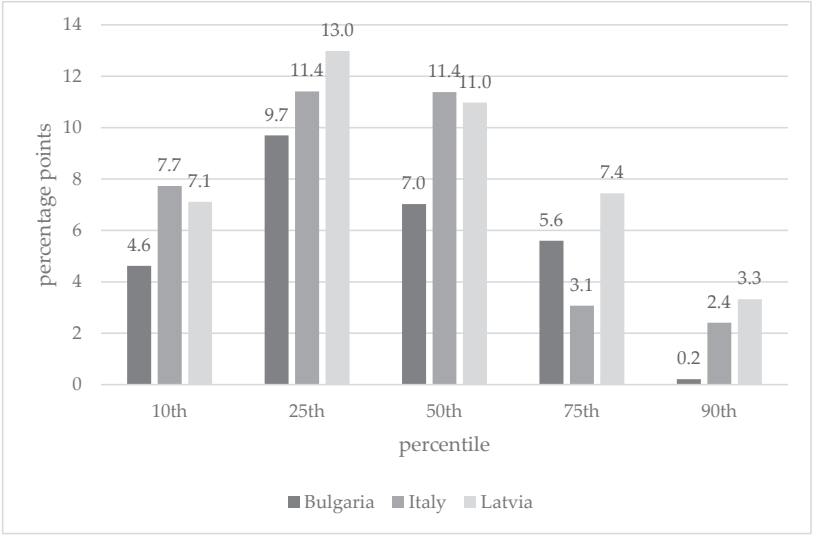


Figure 2. Reduction in the GWG due to an increase in GEMP. Source: own elaboration based on the authors' estimate using E-SES 2018.

The second group of countries consists of Cyprus, France, Malta, Poland and Sweden. In these countries, the estimated coefficient of the cross product of the female and *GEMP* variables is positive and statistically significant in the middle and lower parts of the wage distribution and becomes negative moving up the distribution. This indicates that in these countries, a higher *GEMP* reduces the wage difference between low-paid women and men and increases it for high-paid workers. Among these countries, Cyprus presents by far the largest coefficients at the middle and bottom parts of the distribution. More specifically, the estimated coefficients suggest that, in Cyprus, a one-point increase in the GEMP index, reduces the GWG by 18.7 and 23.2 percentage points at the 10th and 25th percentiles, respectively, compensating considerably for the adjusted GWG, and by 16.7 percentage points (from 20.9% to 4.2%) at the median. In the third group of eight countries (Croatia, Czechia, Denmark, Germany, Hungary, the Netherlands, Portugal and Slovakia), the estimated coefficient of the cross product of the female and *GEMP* variables is positive and statistically significant at the 50th percentile and, in some cases, at the 25th and 75th percentiles. However, the coefficient is negative and significant at the edges of the distribution. Thus, a more gender-balanced composition of management reduces the GWG for workers at the median of the distribution. The countries where the estimated coefficient is highest at the median are Portugal, Slovakia and Hungary. Indeed, a one-point increase in the *GEMP* index is associated with a 19.7, 24.3 and 11.5 percentage point reduction in the adjusted GWG in Portugal, Slovakia and Hungary, respectively. The fourth group includes only Romania, where the estimated coefficient of the cross product of the female and *GEMP* is negative and statistically significant at all percentiles considered. Finally, the fifth group consists of Greece and Spain, where it is difficult to find a clear pattern along the wage distribution.

The results of the UQR including establishment fixed effects, as stated in Equation (6), are summarised in Table 5. Similar to Table 4, the estimated coefficients and their respective standard deviations for the variables female and the cross product of the female and *GEMP* are reported for each country and for the 10th, 25th, 50th, 75th and 90th percentiles of the wage distribution.

Table 4. Summary of unconditional quantile estimates of wages equation.

Country	Variable	q10		q25		q50		q75		q90	
		Coefficient	(s.e.)	Coefficient	(s.e.)	Coefficient	(s.e.)	Coefficient	(s.e.)	Coefficient	(s.e.)
Belgium	female	0.0198 ***	(0.0046)	−0.0462 ***	(0.0043)	−0.0607 ***	(0.0031)	−0.0019	(0.0032)	0.0089 ***	(0.0030)
	GEM	0.0172 ***	(0.0048)	−0.0002	(0.0045)	−0.0121 ***	(0.0035)	−0.0110 ***	(0.0039)	0.0299 ***	(0.0041)
	female × GEM	0.0620 ***	(0.0075)	0.0862 ***	(0.0068)	0.0308 ***	(0.0050)	0.0084	(0.0054)	−0.0077	(0.0055)
Bulgaria	female	−0.0298 ***	(0.0021)	−0.1060 ***	(0.0024)	−0.1967 ***	(0.0025)	−0.2273 ***	(0.0031)	−0.2007 ***	(0.0039)
	GEM	−0.0128 ***	(0.0023)	−0.0004	(0.0025)	0.0345 ***	(0.0027)	0.0228 ***	(0.0036)	0.0628 ***	(0.0046)
	female × GEM	0.0465 ***	(0.0033)	0.1024 ***	(0.0037)	0.0821 ***	(0.0038)	0.0679 ***	(0.0048)	0.0026	(0.0060)
Cyprus	female	−0.1588 ***	(0.0064)	−0.2423 ***	(0.0052)	−0.2339 ***	(0.0065)	−0.0379 ***	(0.0052)	−0.0164 ***	(0.0053)
	GEM	−0.0053	(0.0081)	0.0244 ***	(0.0073)	−0.1025 ***	(0.0098)	0.0451 ***	(0.0091)	0.0548 ***	(0.0105)
	female × GEM	0.1985 ***	(0.0099)	0.2588 ***	(0.0088)	0.1912 ***	(0.0120)	−0.0993 ***	(0.0107)	−0.1646 ***	(0.0119)
Czechia	female	−0.0971 ***	(0.0014)	−0.1501 ***	(0.0010)	−0.1854 ***	(0.0008)	−0.2188 ***	(0.0010)	−0.2282 ***	(0.0016)
	GEM	0.0180 ***	(0.0017)	−0.0435 ***	(0.0012)	−0.0716 ***	(0.0010)	−0.0637 ***	(0.0013)	−0.0098 ***	(0.0024)
	female × GEM	−0.0313 ***	(0.0025)	0.0041 **	(0.0017)	0.0426 ***	(0.0014)	0.0212 ***	(0.0017)	−0.0620 ***	(0.0029)
Germany	female	0.0246 ***	(0.0012)	−0.0307 ***	(0.0006)	−0.1247 ***	(0.0007)	−0.1269 ***	(0.0007)	−0.1514 ***	(0.0010)
	GEM	−0.0218 ***	(0.0016)	−0.0153 ***	(0.0009)	−0.0176 ***	(0.0009)	0.0114 ***	(0.0010)	0.0279 ***	(0.0017)
	female × GEM	−0.0207 ***	(0.0023)	0.0041 ***	(0.0012)	0.0244 ***	(0.0013)	−0.0103 ***	(0.0012)	−0.0597 ***	(0.0019)
Denmark	female	−0.0543 ***	(0.0015)	−0.0859 ***	(0.0010)	−0.1017 ***	(0.0009)	−0.1328 ***	(0.0012)	−0.1639 ***	(0.0020)
	GEM	−0.0080 ***	(0.0015)	−0.0152 ***	(0.0011)	−0.0041 ***	(0.0010)	0.0268 ***	(0.0014)	0.0492 ***	(0.0027)
	female × GEM	−0.0002	(0.0023)	0.0334 ***	(0.0016)	0.0304 ***	(0.0014)	0.0000	(0.0018)	−0.0268 ***	(0.0031)
Estonia	female	−0.0559 ***	(0.0076)	−0.2175 ***	(0.0099)	−0.1847 ***	(0.0055)	−0.2000 ***	(0.0068)	−0.2757 ***	(0.0110)
	GEM	−0.1570 ***	(0.0095)	−0.1536 ***	(0.0118)	−0.1014 ***	(0.0066)	−0.0370 ***	(0.0088)	−0.0581 ***	(0.0151)
	female × GEM	−0.0097	(0.0137)	0.0309 *	(0.0164)	0.0878 ***	(0.0088)	0.0250 **	(0.0111)	0.0540 ***	(0.0180)
Greece	female	−0.0460 ***	(0.0069)	−0.0749 ***	(0.0053)	−0.0583 ***	(0.0049)	−0.1524 ***	(0.0061)	−0.2216 ***	(0.0115)
	GEM	−0.1065 ***	(0.0084)	−0.0576 ***	(0.0062)	0.0397 ***	(0.0060)	−0.0417 ***	(0.0078)	0.0013	(0.0167)
	female × GEM	0.0347 ***	(0.0120)	−0.0066	(0.0090)	−0.0743 ***	(0.0081)	0.0047	(0.0101)	0.0213	(0.0194)
Spain	female	−0.1436 ***	(0.0028)	−0.1684 ***	(0.0021)	−0.1600 ***	(0.0023)	−0.0924 ***	(0.0022)	−0.1093 ***	(0.0027)
	GEM	−0.0259 ***	(0.0026)	0.0268 ***	(0.0021)	0.0247 ***	(0.0023)	0.0291 ***	(0.0023)	−0.0280 ***	(0.0031)
	female × GEM	0.0067	(0.0044)	0.0530 ***	(0.0033)	−0.0003	(0.0034)	−0.0200 ***	(0.0033)	0.0372 ***	(0.0042)
France	female	−0.0917 ***	(0.0008)	−0.0990 ***	(0.0007)	−0.0934 ***	(0.0006)	−0.1520 ***	(0.0008)	−0.1960 ***	(0.0013)
	GEM	−0.0049 ***	(0.0006)	−0.0007	(0.0006)	0.0037 ***	(0.0006)	0.0119 ***	(0.0009)	0.0387 ***	(0.0014)
	female × GEM	0.0270 ***	(0.0012)	0.0190 ***	(0.0011)	−0.0025 ***	(0.0009)	−0.0194 ***	(0.0013)	−0.0316 ***	(0.0020)

Table 4. *Cont.*

Country	Variable	q10		q25		q50		q75		q90	
		Coefficient	(s.e.)	Coefficient	(s.e.)	Coefficient	(s.e.)	Coefficient	(s.e.)	Coefficient	(s.e.)
Croatia	female	−0.1540 ***	(0.0041)	−0.1698 ***	(0.0038)	−0.1521 ***	(0.0035)	−0.1592 ***	(0.0045)	−0.1898 ***	(0.0066)
	GEM	−0.0431 ***	(0.0035)	−0.0130 ***	(0.0036)	−0.0235 ***	(0.0035)	−0.0110 ***	(0.0041)	−0.0182 ***	(0.0061)
	female × GEM	−0.0061	(0.0061)	−0.0172 ***	(0.0055)	0.0229 ***	(0.0051)	0.0108 *	(0.0063)	−0.0095	(0.0091)
Hungary	female	−0.0046 ***	(0.0012)	−0.0250 ***	(0.0013)	−0.1152 ***	(0.0016)	−0.1186 ***	(0.0017)	−0.2366 ***	(0.0033)
	GEM	−0.0058 ***	(0.0014)	−0.0192 ***	(0.0015)	−0.0392 ***	(0.0018)	0.0070 ***	(0.0020)	0.0469 ***	(0.0044)
	female × GEM	−0.0115 ***	(0.0018)	−0.0014	(0.0019)	0.0587 ***	(0.0024)	−0.0160 ***	(0.0026)	−0.0653 ***	(0.0053)
Italy	female	−0.0680 ***	(0.0015)	−0.1207 ***	(0.0013)	−0.1413 ***	(0.0013)	−0.1581 ***	(0.0018)	−0.2423 ***	(0.0039)
	GEM	−0.0102 ***	(0.0017)	−0.0191 ***	(0.0015)	−0.0406 ***	(0.0015)	−0.0094 ***	(0.0025)	−0.1263 ***	(0.0057)
	female × GEM	0.0795 ***	(0.0022)	0.1211 ***	(0.0020)	0.1232 ***	(0.0020)	0.0353 ***	(0.0032)	0.0302 ***	(0.0068)
Latvia	female	−0.1125 ***	(0.0028)	−0.2212 ***	(0.0036)	−0.2022 ***	(0.0025)	−0.2417 ***	(0.0032)	−0.3019 ***	(0.0054)
	GEM	−0.0412 ***	(0.0029)	−0.0784 ***	(0.0040)	−0.0446 ***	(0.0029)	−0.0113 ***	(0.0038)	0.0153 **	(0.0065)
	female × GEM	0.0766 ***	(0.0043)	0.1501 ***	(0.0056)	0.1261 ***	(0.0039)	0.0906 ***	(0.0050)	0.0440 ***	(0.0083)
Malta	female	−0.1033 ***	(0.0071)	−0.0818 ***	(0.0063)	−0.0343 ***	(0.0054)	−0.0271 ***	(0.0065)	−0.1654 ***	(0.0109)
	GEM	−0.0641 ***	(0.0071)	−0.0962 ***	(0.0068)	−0.0046	(0.0061)	−0.0311 ***	(0.0078)	−0.0093	(0.0146)
	female × GEM	0.0540 ***	(0.0107)	0.0095	(0.0097)	−0.0632 ***	(0.0084)	−0.0875 ***	(0.0102)	−0.0054	(0.0170)
Netherlands	female	0.0358 ***	(0.0011)	−0.0646 ***	(0.0012)	−0.0719 ***	(0.0009)	−0.1093 ***	(0.0010)	−0.1245 ***	(0.0016)
	GEM	0.0141 ***	(0.0013)	0.0320 ***	(0.0015)	0.0301 ***	(0.0010)	0.0193 ***	(0.0012)	0.0521 ***	(0.0021)
	female × GEM	−0.0357 ***	(0.0017)	0.0272 ***	(0.0019)	0.0155 ***	(0.0014)	−0.0152 ***	(0.0015)	−0.0303 ***	(0.0025)
Norway	female	−0.0372 ***	(0.0016)	−0.0642 ***	(0.0011)	−0.0893 ***	(0.0009)	−0.1464 ***	(0.0013)	−0.1896 ***	(0.0022)
	GEM	−0.0329 ***	(0.0017)	−0.0254 ***	(0.0011)	−0.0040 ***	(0.0010)	0.0130 ***	(0.0015)	0.0079 ***	(0.0028)
	female × GEM	0.0224 ***	(0.0025)	0.0197 ***	(0.0017)	0.0130 ***	(0.0014)	0.0157 ***	(0.0020)	0.0047	(0.0033)
Poland	female	−0.0505 ***	(0.0009)	−0.1285 ***	(0.0008)	−0.1795 ***	(0.0007)	−0.2078 ***	(0.0010)	−0.1988 ***	(0.0014)
	GEM	0.0027 ***	(0.0009)	−0.0272 ***	(0.0008)	−0.0423 ***	(0.0009)	−0.0047 ***	(0.0011)	0.0384 ***	(0.0015)
	female × GEM	0.0181 ***	(0.0014)	0.0346 ***	(0.0012)	0.0448 ***	(0.0011)	0.0129 ***	(0.0016)	−0.0378 ***	(0.0022)
Portugal	female	−0.0794 ***	(0.0020)	−0.0988 ***	(0.0023)	−0.1936 ***	(0.0043)	−0.1960 ***	(0.0053)	−0.1556 ***	(0.0054)
	GEM	−0.0668 ***	(0.0020)	−0.0621 ***	(0.0023)	−0.1433 ***	(0.0048)	0.0509 ***	(0.0055)	0.0878 ***	(0.0063)
	female × GEM	0.0050	(0.0031)	0.0588 ***	(0.0034)	0.1139 ***	(0.0067)	−0.0469 ***	(0.0079)	−0.1104 ***	(0.0084)
Romania	female	−0.0016	(0.0014)	−0.0842 ***	(0.0018)	−0.1067 ***	(0.0016)	−0.1076 ***	(0.0019)	−0.1320 ***	(0.0026)
	GEM	−0.0271 ***	(0.0014)	−0.0020	(0.0017)	−0.0089 ***	(0.0016)	0.0303 ***	(0.0017)	0.0789 ***	(0.0023)
	female × GEM	−0.0141 ***	(0.0018)	−0.0347 ***	(0.0023)	−0.0543 ***	(0.0021)	−0.0450 ***	(0.0024)	−0.0505 ***	(0.0033)

Table 4. *Cont.*

Country	Variable	q10		q25		q50		q75		q90	
		Coefficient	(s.e.)	Coefficient	(s.e.)	Coefficient	(s.e.)	Coefficient	(s.e.)	Coefficient	(s.e.)
Sweden	female	−0.0157 ***	(0.0009)	−0.0235 ***	(0.0007)	−0.0475 ***	(0.0008)	−0.0642 ***	(0.0012)	−0.1359 ***	(0.0020)
	GEM	0.0107 ***	(0.0009)	0.0059 ***	(0.0007)	−0.0031 ***	(0.0008)	0.0294 ***	(0.0014)	0.0262 ***	(0.0025)
	female × GEM	0.0057 ***	(0.0014)	0.0054 ***	(0.0011)	0.0025 **	(0.0012)	−0.0358 ***	(0.0018)	−0.0075 **	(0.0031)
Slovakia	female	−0.0751 ***	(0.0021)	−0.1701 ***	(0.0016)	−0.2520 ***	(0.0015)	−0.2625 ***	(0.0019)	−0.2204 ***	(0.0028)
	GEM	0.0051 **	(0.0024)	−0.0324 ***	(0.0018)	−0.0703 ***	(0.0019)	−0.0554 ***	(0.0025)	0.0441 ***	(0.0040)
	female × GEM	−0.0074 **	(0.0035)	0.0494 ***	(0.0026)	0.0880 ***	(0.0025)	0.0321 ***	(0.0031)	−0.1067 ***	(0.0049)

Notes: Sampling weights are used in all estimations. Robust standard errors (s.e.) between parentheses. Statistically significance indicated by *** 1% level, ** 5% level and * 10% level. Source: own elaboration based on the authors' estimate using E-SES 2018.

Table 5. Summary of unconditional quantile estimates of wage equation with establishment fixed effects.

Country	Variable	q10		q25		q50		q75		q90	
		Coefficient	(s.e.)	Coefficient	(s.e.)	Coefficient	(s.e.)	Coefficient	(s.e.)	Coefficient	(s.e.)
Belgium	female	0.0359	(0.0348)	−0.0241	(0.0249)	−0.0675 ***	(0.0155)	−0.0258	(0.0170)	0.0030	(0.0154)
	female × GEM	0.0413	(0.0567)	0.0555	(0.0419)	0.0509 *	(0.0291)	0.0449	(0.0323)	−0.0030	(0.0328)
Bulgaria	female	−0.0107	(0.0069)	−0.0559 ***	(0.0077)	−0.1174 ***	(0.0077)	−0.1473 ***	(0.0103)	−0.1835 ***	(0.0148)
	female × GEM	0.0208 *	(0.0107)	0.0437 ***	(0.0117)	0.0242 **	(0.0121)	0.0177	(0.0161)	0.0218	(0.0232)
Cyprus	female	−0.1214 ***	(0.0247)	−0.1748 ***	(0.0219)	−0.1707 ***	(0.0259)	−0.0282	(0.0216)	−0.0069	(0.0248)
	female × GEM	0.1667 ***	(0.0378)	0.1912 ***	(0.0361)	0.1198 **	(0.0583)	−0.1261 ***	(0.0481)	−0.1667 **	(0.0649)
Czechia	female	−0.0876 ***	(0.0049)	−0.1202 ***	(0.0031)	−0.1522 ***	(0.0024)	−0.1825 ***	(0.0030)	−0.1999 ***	(0.0052)
	female × GEM	−0.0061	(0.0082)	0.0198 ***	(0.0052)	0.0527 ***	(0.0039)	0.0337 ***	(0.0048)	−0.0644 ***	(0.0087)
Germany	female	0.0006	(0.0123)	−0.0418 ***	(0.0072)	−0.1411 ***	(0.0074)	−0.1513 ***	(0.0070)	−0.2347 ***	(0.0128)
	female × GEM	0.0005	(0.0212)	0.0069	(0.0123)	0.0503 ***	(0.0125)	0.0187	(0.0120)	0.0326	(0.0219)
Denmark	female	−0.0309 ***	(0.0020)	−0.0565 ***	(0.0014)	−0.0818 ***	(0.0012)	−0.1266 ***	(0.0016)	−0.1680 ***	(0.0029)
	female × GEM	−0.0005	(0.0031)	0.0162 ***	(0.0022)	0.0175 ***	(0.0018)	0.0033	(0.0026)	−0.0246 ***	(0.0047)
Estonia	female	−0.0303	(0.0239)	−0.1499 ***	(0.0284)	−0.1138 ***	(0.0167)	−0.1013 ***	(0.0203)	−0.1977 ***	(0.0287)
	female × GEM	0.0228	(0.0388)	0.0673	(0.0455)	0.0518 **	(0.0260)	−0.0366	(0.0312)	0.0473	(0.0472)
Greece	female	−0.0043	(0.0421)	−0.0791 *	(0.0444)	−0.0481	(0.0403)	−0.1314 ***	(0.0483)	−0.1605	(0.1014)
	female × GEM	0.0011	(0.0750)	0.0199	(0.0716)	−0.0662	(0.0622)	−0.0202	(0.0747)	−0.0452	(0.1578)

Table 5. *Cont.*

Country	Variable	q10		q25		q50		q75		q90	
		Coefficient	(s.e.)	Coefficient	(s.e.)	Coefficient	(s.e.)	Coefficient	(s.e.)	Coefficient	(s.e.)
Spain	female	−0.1275 ***	(0.0246)	−0.1359 ***	(0.0251)	−0.1516 ***	(0.0257)	−0.0984 ***	(0.0233)	−0.1411 ***	(0.0304)
	female × GEM	0.0419	(0.0376)	0.0528	(0.0358)	0.0228	(0.0351)	−0.0024	(0.0318)	0.0776 *	(0.0433)
France	female	−0.0683 ***	(0.0132)	−0.0861 ***	(0.0115)	−0.0877 ***	(0.0086)	−0.1342 ***	(0.0119)	−0.1925 ***	(0.0152)
	female × GEM	0.0228	(0.0181)	0.0360 **	(0.0159)	−0.0048	(0.0121)	−0.0486 ***	(0.0168)	−0.0647 ***	(0.0232)
Croatia	female	−0.0614 ***	(0.0142)	−0.0678 ***	(0.0127)	−0.0748 ***	(0.0135)	−0.1201 ***	(0.0188)	−0.1636 ***	(0.0271)
	female × GEM	−0.0466 **	(0.0223)	−0.0625 ***	(0.0196)	−0.0127	(0.0209)	−0.0041	(0.0284)	−0.0474	(0.0409)
Hungary	female	−0.0418 ***	(0.0050)	−0.0230 ***	(0.0043)	−0.0720 ***	(0.0053)	−0.0918 ***	(0.0055)	−0.1986 ***	(0.0109)
	female × GEM	0.0307 ***	(0.0071)	0.0122 **	(0.0062)	0.0340 ***	(0.0075)	−0.0195 **	(0.0079)	−0.0511 ***	(0.0161)
Italy	female	−0.0389 ***	(0.0123)	−0.0899 ***	(0.0124)	−0.1185 ***	(0.0115)	−0.1327 ***	(0.0177)	−0.1929 ***	(0.0389)
	female × GEM	0.0362 *	(0.0189)	0.0755 ***	(0.0179)	0.0954 ***	(0.0176)	−0.0247	(0.0297)	−0.0656	(0.0654)
Latvia	female	−0.0273 **	(0.0112)	−0.0550 ***	(0.0136)	−0.0722 ***	(0.0095)	−0.1282 ***	(0.0126)	−0.1648 ***	(0.0223)
	female × GEM	−0.0126	(0.0186)	−0.0028	(0.0215)	−0.0048	(0.0151)	−0.0340 *	(0.0198)	−0.1058 ***	(0.0351)
Malta	female	−0.1428 ***	(0.0314)	−0.1485 ***	(0.0265)	−0.0504 **	(0.0238)	−0.0705 **	(0.0292)	−0.1690 ***	(0.0492)
	female × GEM	0.1231 ***	(0.0434)	0.1267 ***	(0.0379)	−0.0262	(0.0335)	0.0015	(0.0427)	0.0087	(0.0750)
Netherlands	female	0.0373 ***	(0.0144)	−0.0352 ***	(0.0112)	−0.0675 ***	(0.0072)	−0.1258 ***	(0.0083)	−0.1429 ***	(0.0117)
	female × GEM	−0.0507 **	(0.0226)	0.0133	(0.0186)	0.0183	(0.0115)	0.0203	(0.0130)	0.0199	(0.0182)
Norway	female	−0.0014	(0.0019)	−0.0288 ***	(0.0012)	−0.0545 ***	(0.0010)	−0.1058 ***	(0.0014)	−0.1690 ***	(0.0024)
	female × GEM	−0.0044	(0.0028)	−0.0039 **	(0.0018)	−0.0161 ***	(0.0016)	−0.0221 ***	(0.0023)	−0.0194 ***	(0.0038)
Poland	female	−0.0370 ***	(0.0030)	−0.0829 ***	(0.0027)	−0.1121 ***	(0.0026)	−0.1598 ***	(0.0035)	−0.1549 ***	(0.0050)
	female × GEM	0.0090 *	(0.0047)	0.0136 ***	(0.0042)	0.0130 ***	(0.0040)	−0.0038	(0.0055)	−0.0760 ***	(0.0083)
Portugal	female	−0.0455 ***	(0.0141)	−0.0616 ***	(0.0153)	−0.1804 ***	(0.0265)	−0.2299 ***	(0.0334)	−0.1936 ***	(0.0360)
	female × GEM	−0.0076	(0.0296)	0.0415	(0.0265)	0.1362 ***	(0.0411)	0.0328	(0.0486)	−0.0240	(0.0551)
Romania	female	−0.0347 ***	(0.0053)	−0.0785 ***	(0.0090)	−0.0902 ***	(0.0094)	−0.1206 ***	(0.0115)	−0.1232 ***	(0.0151)
	female × GEM	0.0212 ***	(0.0072)	−0.0171	(0.0115)	−0.0499 ***	(0.0118)	−0.0156	(0.0141)	−0.0576 ***	(0.0194)
Sweden	female	−0.0052	(0.0045)	−0.0144 ***	(0.0037)	−0.0496 ***	(0.0039)	−0.0821 ***	(0.0059)	−0.1606 ***	(0.0112)
	female × GEM	−0.0036	(0.0069)	0.0012	(0.0057)	0.0155 ***	(0.0060)	0.0027	(0.0090)	0.0276	(0.0174)
Slovakia	female	−0.0587 ***	(0.0068)	−0.0970 ***	(0.0048)	−0.1481 ***	(0.0048)	−0.1664 ***	(0.0061)	−0.1670 ***	(0.0112)
	female × GEM	−0.0054	(0.0109)	0.0227 ***	(0.0080)	0.0381 ***	(0.0080)	−0.0111	(0.0104)	−0.1354 ***	(0.0198)

Notes: Sampling weights are used in all estimations. Robust standard errors (s.e.) between parentheses. Statistically significance indicated by *** 1% level, ** 5% level and * 10% level. Source: own elaboration based on the authors' estimate using E-SES 2018.

In line with the results previously presented in Table 4, the estimated coefficient for the variable female is lowest in absolute value at the bottom of the wage distribution and increases in magnitude as we move up along the wage distribution. This evidence of the presence of the glass ceiling for women in the job market is found in all countries except for Belgium and Cyprus. The countries where the difference in the GWG between the 90th and 10th percentile of the wage distribution is larger are Germany, the Netherlands and Bulgaria. The difference is by almost 21 percentage points in Germany, 17.1 percentage points in the Netherlands, and 15.7 percentage points in Bulgaria.

Regarding the effect of *GEMP* on the GWG, the results are similar to those without establishment fixed effects. Thus, gender equality in management reduces the GWG in the middle and lower part of the wage distribution, while a greater gender equality is associated with a higher GWG at the top of the distribution. This may be attributed to the standardisation of lower-wage jobs, typically established through collective agreements, while higher wages are usually negotiated based on the perceived value contributed by workers, with different gender biases playing a role. Furthermore, it is usual for higher wages to be linked to perceived wage bonuses that reward greater effort in terms of time, travel and the volume of projects undertaken, which women often do not have because they are engaged in care work (De la Rica et al. 2015; Christofides et al. 2013).

However, there is a significant increase in the number of cases where the estimated coefficient is not statistically significant. At the 10th, 25th and 50th percentiles the number of countries where the estimated coefficient of the cross product of the female and *GEMP* variables is positive and statistically significant is 7, 10 and 13, respectively, while the coefficient is negative in only two countries at each indicated percentile. At the 75th percentile, the estimated coefficient is negative (positive) in five countries (one country). At the 90th percentile, the estimated coefficient is negative (positive) in 10 countries (1 country). The countries can be now divided into four groups. The first group consists of countries where greater gender equality reduces the wage gap between women and men at the lower and medium parts of the wage distribution and increases it between high-paid women and men. These countries are Bulgaria, Cyprus, Czechia, Denmark, France, Hungary, Italy, Malta, Poland, Romania and Slovakia. The second group consists of Belgium, Estonia, Germany, Portugal and Sweden. In these countries, the estimated coefficient of the cross product of the female and *GEMP* is positive and statistically significant at the median and not statistically significant in the rest of the wage distribution. Among these, the estimated coefficient is larger in Portugal. Its value indicates that a one-point increase in the *GEMP* index reduces the GWG at the 50th percentile by 12.2 percentage points, from 16.5% to 4.3%. For Belgium, Germany and Estonia, the estimated coefficient is rather similar (around 0.05), indicating that a one-point increase in the *GEMP* index reduces the GWG at the median by 4.8 percentage points in Belgium (from 6.5% to 1.7%), 4.5 percentage points in Germany (from 13.2% to 8.6%) and 4.7 percentage points in Estonia (from (10.8% to 6%). In the third group of countries, the estimated coefficient of the cross product of the female and *GEMP* is negative and statistically significant at one or more points of the wage distribution. This group consists of Croatia, Latvia, Norway and the Netherlands. Finally, in Spain and Greece, the results suggest that gender equality in management has no effect on the GWG at any point of the wage distribution.

6. Discussion

Encouraging women to promote to leadership positions is one of the mechanisms proposed in the literature and policies to reduce the GWG. Involving more women in decision making would prioritise gender equality, facilitating the implementation of policies and practices, like family-friendly policies, enabling women to compete on an equal term. Previous studies have showed that countries with less generous gender policies have a lower pay gap at the top of the wage distribution and a larger gap at the bottom, suggesting that the positive effect of family-friendly policies dominates at the bottom of the distribution (Arulampalam et al. 2007). Such policies do indeed correlate with the

duration of employment breaks taken by women after the birth of their first child. This correlation is significant when female employees are specifically targeted for promotion, through initiatives such as mentoring programmes, which lead to a faster return to work (Bächmann et al. 2020).

The legal framework of the European Union acts as a keystone for closing the gender pay gap that reveals possible gender stereotypes and bias, contributing to the endurance of the other cause of the GWG, like sectoral and vertical segregation—the glass ceiling (Leythienne and Pérez-Julián 2021). The European Union itself has already introduced gender quotas into their legislation to reduce the glass ceiling. Although certain studies indicate that quotas have not achieved their purpose, the political and social obligation to report on board composition and pay gaps generates public image pressure to improve gender equality (Theodoropoulos et al. 2022).

The absence or presence of a structured childcare system in the countries could cause women managers to be less or more attached to the labour market, especially at the bottom of the wage distribution. This characteristic along with social and cultural biases may constitute a significant obstacle to equality between men and women in terms of wages and participation in managerial position, particularly for the high hierarchical levels (Sccichiano 2014).

Cultural changes could have led more highly and technical-educated women to choose stereotypically male work in order to pursue goals of increased status and pay (Busch 2020). In this sense, the challenges are associated with the continuous requirement for up-skilling and reskilling as an adaptive reaction to technological changes that will affect the labour market and the new jobs (Cramarenco et al. 2023). Artificial Intelligence is expected to generate disruptive transformations in the labour market, affecting job requirements, work procedures, task design, and assessment techniques. These transformations will also impact gender equality. Therefore, closing the gender gaps—both the glass ceiling and the GWG—depends on society's continued support of women's training in STEM and AI and ensuring that today's gender stereotypes do not become internalised in future technology systems (Collett et al. 2022).

7. Conclusions

Our work includes a gender-neutral diversity perspective to contribute to a deeper understanding of the dynamics through which gender diversity might reduce the GWG. In addition to examining only the proportion of women in leadership positions, this approach broadens the scope of the evaluation to include the impact of men's entry into women-led organisations. Moreover, studying its effects throughout the income distribution offers additional valuable perspectives.

Our findings indicate the relationship between gender equality in management positions and wages yields mixed results. In 10 of the 22 countries, higher gender equality in management corresponds to lower wages, while in 6 of the 22 countries, there is evidence that a greater level of gender equality in management is associated with increased wages. Furthermore, the results show that gender equality in management has a predominantly positive effect in the upper part of the wage distribution, and a negative effect at the middle and lower parts of the distribution.

Regarding the relationship between gender equality in managerial positions and the GWG, the results are also mixed, but the cases where a higher gender equality reduces the GWG predominate. Also, the estimation results of the UQR indicate that gender equality in managerial positions reduces the GWG mainly in the middle and lower parts of the wage distribution, in accordance with the results obtained by Santero-Sánchez and Castro Núñez (2022) for Spain and Huffman et al. (2017) for Germany.

There are several possible causes for these findings. Firstly, the current number of women in managerial positions might not be adequate to bring about the required transformations. Specifically, their increased representation may not be enough to provide the power necessary to change the sexist structures that generate the GWG. Secondly, if

gender equality in management is mainly achieved in feminised sectors, the lower salaries in these sectors would inhibit a reduction in the GWG at the global level.

Moreover, it is possible that gender diversity in corporate management requires a time lag for changes to become visible within organisations. It would be interesting to test whether the influence of gender diversity affects wages with a time lapse, in the sense that the effect could be observed in later years. Therefore, it is needed to have a panel database that would allow us to analyse changes in the variables over time in the same set of companies.

On the other hand, these results represent a policy dilemma, as promoting gender equality in management improves the situation for the majority of women who are concentrated at the bottom of the wage distribution, while disadvantaging a few at the top (Huffman et al. 2017).

Author Contributions: Conceptualization and literature review, L.d.C.R., V.M.B. and R.S.-S.; methodology, L.d.C.R., V.M.B. and R.S.-S.; database treatment and statistical analysis, L.d.C.R., V.M.B. and R.S.-S.; writing—original draft preparation, L.d.C.R., V.M.B. and R.S.-S.; writing—review and editing, L.d.C.R., V.M.B. and R.S.-S. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data used in this study are available for scientific purposes only upon request to entities recognized as research entities by Eurostat. More information can be found at https://ec.europa.eu/eurostat/web/microdata/structure-of-earnings-survey accessed on 27 March 2023.

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

Notes

1 The formula used to calculate the index of gender equality in managerial positions is based on the gender wage gap measure used in the construction of the Gender Equality Index by the European Institute for Gender Equality (European Institute for Gender Equality 2017).

2 Equation (6) is estimated using the STATA module rifghde by Ríos-Ávila (2020). The command uses a two-step procedure. First, RIF are estimated as in Equation (4), and second, OLS with one high-dimensional fixed effect (the establishment fixed effect) is applied. The module is part of the STATA package RIF available at: https://github.com/friosavila/stpackages/tree/main/rif (accessed on 27 March 2023).

3 International Standard Classification of Occupations, year 2008.

4 We do not present the results for the rest of control variables due to space limitations. The overall results are available upon request on the authors.

References

Abraham, Mabel. 2017. Pay Formalization Revisited: Considering the Effects of Manager Gender and Discretion on Closing the Gender Wage Gap. *Academy of Management Journal* 60: 29–54. [CrossRef]

Akerlof, George A., and Rachel E. Kranton. 2000. Economics and identity. *The Quarterly Journal of Economics* 115: 715–53. [CrossRef]

Albrecht, James, Anders Bjorklund, and Susan Vroman. 2003. Is there a glass ceiling in Sweden? *Journal of Labor Economics* 21: 145–77. [CrossRef]

Arulampalam, Wiji, Alison L. Booth, and Mark L. Bryan. 2007. Is There a Glass Ceiling over Europe? Exploring the Gender Pay Gap across the Wage Distribution. *ILR Review* 60: 163–86. [CrossRef]

Bächmann, Ann-Christin, Corinna Frodermann, and Dana Müller. 2020. Does the Firm Make the Difference? The Influence of Organizational Family-Friendly Arrangements on the Duration of Employment Interruptions after Childbirth. *European Sociological Review* 36: 798–813. [CrossRef]

Beaman, Lori, Esther Duflo, Rohini Pande, and Petia Topalova. 2012. Female Leadership Raises Aspirations and Educational Attainment for Girls. *Science* 335: 582–86. [CrossRef] [PubMed]

Beaman, Lori, Raghabendra Chattopadhyay, Esther Duflo, Rohini Pande, and Petia Topalova. 2009. Powerful Women: Does Exposure Reduce Bias? *The Quarterly Journal of Economics* 124: 1497–540. [CrossRef]

Becker, Gary S. 1957. *The Economics of Discrimination*. Chicago: University of Chicago Press, 137p.

Becker, Gary S. 1962. Investment in human capital: A theoretical analysis. *Journal of Political Economy* 70: 9–49. [CrossRef]

Bell, Linda A. 2005. Women-led firms and the gender gap in top executive jobs. *IZA Discussion Paper 1689*. [CrossRef]

Bertrand, Marianne, Claudia Goldin, and Lawrence F. Katz. 2010. Dynamics of the Gender Gap for Young Professionals in the Financial and Corporate Sectors. *American Economic Journal Applied Economics* 2: 228–55. [CrossRef]
Bertrand, Marianne, Sandra E. Black, Sissel Jensen, and Adriana Lleras-Muney. 2019. Breaking the Glass Ceiling? The Effect of Board Quotas on Female Labour Market Outcomes in Norway. *The Review of Economic Studies* 86: 191–239. [CrossRef]
Blau, Francine D., and Lawrence M. Kahn. 2017. The gender wage gap: Extent, trends, and explanations. *Journal of Economic Literature* 55: 789–865. [CrossRef]
Busch, Felix. 2020. Gender Segregation, Occupational Sorting, and Growth of Wage Disparities Between Women. *Demography* 57: 1063–1088. [CrossRef]
Cardoso, Ana R., and Rudolf Winter-Ebmer. 2010. Female-led firms and gender wage policies. *ILR Review* 64: 143–63. [CrossRef]
Casado-Díaz, José M., Oana Driha, and Hipólito Simón. 2020. The Gender Wage Gap in Hospitality: New Evidence from Spain. *Cornell Hospitality Quarterly* 63: 399–417. [CrossRef]
Christofides, Louis N., Alexandros Polycarpou, and Konstantinos Vrachimis. 2013. Gender wage gaps, 'sticky floors' and glass ceilings in Europe. *Labor Economics* 21: 86–102. [CrossRef]
Chzhen, Yekaterina, and Karen Mumford. 2011. Gender gaps across the earnings distribution for full-time employees in Britain: Allowing for sample selection. *Labour Economics* 18: 837–44. [CrossRef]
Cohen, Philip N., and Matt L. Huffman. 2007. Working for the Woman? Female Managers and the Gender Wage Gap. *American Sociological Review* 72: 681–704. [CrossRef]
Collett, Clementine, Gina Neff, and Livia Gouvea. 2022. *Los Efectos de la IA en la vida Laboral de las Mujeres*. Paris: UNESCO, Organisation for Economic Co-operation and Development and Inter-American Development Bank. [CrossRef]
Conde-Ruiz, J. Ignacio, and Ignacio Marra de Artíñano. 2016. Gender gaps in the Spanish labor market. In *Estudios sobre la Economía Española—2016/32*. Madrid: FEDEA.
Cotter, David A., Joan M. Hermsen, Seth Ovadia, and Reeve Vanneman. 2001. The Glass Ceiling Effect. *Social Forces* 80: 655–81. [CrossRef]
Cramarenco, Romana Emilia, Monica Ioana Burcă-Voicu, and Dan-Cristian Dabija. 2023. The impact of artificial intelligence (AI) on employees' skills and well-being in global labor markets: A systematic review. *Oeconomia Copernicana* 14: 731–67. [CrossRef]
Dalvit, Nicolò, Aseem Patel, and Joanne Tan. 2021. Intra-Firm Hierarchies and Gender Gaps. *Labour Economics* 77: 102029. [CrossRef]
De la Rica, Sara, Juan J. Dolado, and Raquel Vegas. 2015. Gender Gaps in Performance Pay: New Evidence from Spain. *Annals of Economics and Statistics* 117/118: 41–59. [CrossRef]
Derks, Belle, Naomi Ellemers, Colette van Laar, and Kim de Groot. 2011. Do sexist organizational cultures create the Queen Bee? *British Journal of Social Psychology* 50: 519–35. [CrossRef]
Eagly, Alice H., and Steven J. Karau. 2002. Role Congruity Theory of Prejudice Toward Female Leaders. *Psychological Review* 109: 573–98. [CrossRef]
England, Paula. 1992. *Comparable Worth: Theories and Evidence*. New York: Routledge. [CrossRef]
Eurofound and EIGE. 2021. *Upward Convergence in Gender Equality: How Close Is the Union of Equality?* Luxembourg: Publications Office of the European Union.
European Commission. 2021. *Report on Equality between Women and Men in the EU 2021*. Luxembourg: Publications Office of the European Union. Available online: https://ec.europa.eu/info/sites/default/files/aid_development_cooperation_fundamental_rights/annual_report_ge_2021_en.pdf (accessed on 27 March 2023).
European Institute for Gender Equality. 2017. Gender Equality Index 2017: Methodological Report. Available online: https://eige.europa.eu/publications-resources/publications/gender-equality-index-2017-methodological-report?language_content_entity=en (accessed on 10 May 2023).
Eurostat. 2020. 8 March 2020: International Women's Day. 40/2020–6 March 2020. Available online: https://ec.europa.eu/eurostat/documents/2995521/10474926/3-06032020-AP-EN.pdf/763901be-81b7-ecd6-534e-8a2b83e82934 (accessed on 10 May 2023).
Eurostat. 2023. Data Browser, Gender Pay Gap in Unadjusted Form. Available online: https://ec.europa.eu/eurostat/databrowser/view/sdg_05_20/default/table?lang=en (accessed on 10 May 2023).
Firpo, Sergio, Nicole M. Fortin, and Thomas Lemieux. 2009. Unconditional quantile regressions. *Econometrica* 77: 953–73. Available online: http://www.jstor.org/stable/40263848 (accessed on 27 March 2023).
Flabbi, Luca, Mario Macis, Andrea Moro, and Fabiano Schivardi. 2019. Do female executives make a difference? The impact of female leadership on gender gaps and firm performance. *The Economic Journal* 129: 2390–423. [CrossRef]
Fortin, Nicole M. 2005. Gender role attitudes and the labour-market outcomes of women across OECD countries. *Oxford Review of Economic Policy* 21: 416–38. [CrossRef]
Gharehgozli, Orkideh, and Vidya Atal. 2021. Gender Wage Gap: Within and Across Borders. *Studies in Business and Economics* 16: 74–83. [CrossRef]
Goldin, Claudia. 2014. A grand gender convergence: Its last chapter. *American Economic Review* 104: 1091–119. [CrossRef]
Groshen, Erica L. 1991. The structure of the female/male wage differential: Is it who you are, what you do, or where you work? *The Journal of Human Resources* 26: 457–73. [CrossRef]
Hara, Hiromi. 2018. The gender wage gap across the wage distribution in Japan: Within- and between-establishment effects. *Labour Economics* 53: 213–29. [CrossRef]
Hensvik, Lena E. 2014. Manager impartiality: Worker–firm matching and the gender wage gap. *ILR Review* 67: 395–421. [CrossRef]

Hirsch, Boris. 2013. The impact of female managers on the gender pay gap: Evidence from linked employer–employee data for Germany. *Economics Letters* 119: 348–50. [CrossRef]
Huddy, Leoni, and Nayda Terkildsen. 1993. Gender Stereotypes and the Perception of Male and Female Candidates. *American Journal of Political Science* 37: 119–47. [CrossRef]
Huffman, Matt L. 2004. Gender Inequality Across Local Wage Hierarchies. *Work and Occupations* 31: 323–44. [CrossRef]
Huffman, Matt L. 2016. Does the presence of women in management impact gender inequality? In *Handbook on Well-Being of Working Women*. Edited by Mary L. Connerley and Jiyun Wu. Dordrecht: Springer, pp. 175–87. [CrossRef]
Huffman, Matt L., Joe King, and Malte Reichelt. 2017. Equality for Whom? Organizational Policies and the Gender Gap across the German Earnings Distribution. *ILR Review* 70: 16–41. [CrossRef]
Hultin, Mia, and Ryszard Szulkin. 2003. Mechanisms of inequality: Unequal access to organizational power and the gender wage gap. *European Sociological Review* 19: 143–59. [CrossRef]
International Labour Organization. 2019. Women in Business and Management. The Business Case of Change. International Labor Office, Geneva, Switzerland. Available online: https://www.ilo.org/wcmsp5/groups/public/---dgreports/---dcomm/---publ/documents/publication/wcms_700953.pdf (accessed on 10 May 2023).
Javdani, Mohsen. 2015. Glass ceilings or glass doors? The role of firms in male-female wage disparities. *The Canadian Journal of Economics* 48: 529–60. [CrossRef]
Kalogeraki, Olga, and Dimitrios Georgakakis. 2021. Friend or foe? CEO gender, political ideology, and gender-pay disparities in executive compensation. *Long Range Planning* 55: 102126. [CrossRef]
Killewald, Alexandra, and Jonathan M. Bearak. 2014. Is the motherhood penalty larger for low-wage women? A comment on quantile regression. *American Sociological Review* 79: 350–57. [CrossRef]
Kunze, Astrid, and Amalia Miller. 2017. Women helping women? Evidence from private sector data on workplace hierarchies. *Review of Economics and Statistics* 99: 769–75. [CrossRef]
Leythienne, Denis, and Marina Pérez-Julián. 2021. *Gender Pay Gaps in the European Union-A Statistical Analysis*. Luxembourg: Publications Office of the European Union. [CrossRef]
Magda, Igan, and Ewa Cukrowska-Torzewska. 2019. Do Women Managers Lower Gender Pay Gaps? Evidence from Public and Private Firms. *Feminist Economics* 25: 185–210. [CrossRef]
Mateos de Cabo, Ruth, Ricardo Gimeno, and Lorenzo Escot. 2011. Disentangling discrimination on Spanish boards of directors. *Corporate Governance: An International Review* 19: 77–95. [CrossRef]
Matsa, David A., and Amalia R. Miller. 2011. Chipping away at the glass ceiling: Gender spillovers in corporate leadership. *American Economic Review* 101: 635–39. [CrossRef]
Mincer, Jacob A. 1974. Schooling and Earnings. In *Schooling, Experience, and Earnings*. Cambridge, MA: NBER, pp. 41–63.
Mincer, Jacob, and Solomon Polachek. 1977. An exchange: The theory of human capital and the earnings of women: Women's earnings reexamined. *The Journal of Human Resources* 13: 118–34. [CrossRef]
Phelps, Edmund S. 1972. The statistical theory of racism and sexism. *The American Economic Review* 62: 659–61.
Ríos-Ávila, Fernando. 2020. Recentered influence functions (RIFs) in Stata: RIF regression and RIF decomposition. *The Stata Journal* 20: 51–94. [CrossRef]
Rubery, Jill, Damian Grimshaw, and Hugo Figueiredo. 2005. How to close the gender pay gap in Europe: Towards the gender mainstreaming of pay policy. *Industrial Relations Journal* 36: 184–213. [CrossRef]
Santero-Sánchez, Rosa, and Belén Castro Núñez. 2022. Pursuing equal pay for equal work: Gender diversity in management positions and the gender pay gap throughout the wage distribution. *BRQ Business Research Quarterly* 0: 1–15. [CrossRef]
Sccichiano, Sergio. 2014. The gender wage gap among Spanish managers. *International Journal of Manpower* 35: 327–44. [CrossRef]
Sorensen, Elaine. 1989. Measuring pay disparity between typically female occupations and other jobs: A bivariate selectivity approach. *ILR Review* 42: 624–39. [CrossRef]
Tajfel, Henri, and John C. Turner. 1979. An integrative theory of intergroup conflict. In *The Social Psychology of Intergroup Relations*. Edited by W. G. Austin and S. Worchel. Monterey, CA: Brooks/Cole, pp. 33–37.
Theodoropoulos, Nikolaos, John Forth, and Alex Bryson. 2022. Are Women Doing It for Themselves? Female Managers and the Gender Wage Gap. *Oxford Bulletin of Economics and Statistics* 84: 1329–55. [CrossRef]
Tomaskovic-Devey, Donald. 1993. *Gender and Racial Inequality at Work: The Causes and Consequences of Job Segregation*. Ithaca, NY: ILR Press.
van Hek, Margriet, and Tanja van der Lippe. 2019. Are Female Managers Agents of Change or Cogs in the Machine? An Assessment with Three-Level Manager–Employee Linked Data. *European Sociological Review* 35: 316–31. [CrossRef]
Vega, Pedro J., Rosa Santero, R. Belén Castro, and Nuria Gómez. 2016. Participación femenina en puestos directivos y desigualdad salarial. Un análisis en el mercado laboral español [Women Directors, Critical Mass and Gender Wage Gap. Evidence for Spain]. *Estudios de Economía Aplicada* 34: 155–77.
Zimmermann, Florian. 2022. Managing the Gender Wage Gap—How Female Managers Influence the Gender Wage Gap among Workers. *European Sociological Review* 38: 355–70. [CrossRef]

economies

MDPI

Article

Unraveling the Roots of Income Polarization in Europe: A Divided Continent

Michele Fabiani

Department of Political Science, Communication and International Relations, University of Macerata, Via Don Minzoni 22/A, 62100 Macerata, Italy; m.fabiani9@unimc.it

Abstract: The issue of polarization, as opposed to inequality, has been little explored in European countries. In this paper, using data provided by the Luxembourg Income Studies Database, we look at the trend of income polarization in 12 European countries, the only ones available with two comparable years, using the relative distribution method. The results clearly show a trend toward polarization in almost the cases analyzed, with a concentrated prevalence in the lower tail of the distribution, thus observing a worsening in the distribution. Next, we look at drivers that may have contributed to these changes, using the RIF-regression method. It is interesting to observe how these characteristics are in many cases common across all countries: the occupational sector, level of education and area of residence have the same impact, albeit with different intensities, in all countries. This suggests the possibility of coordinated intervention across these nations, acting on the same variables for all of them.

Keywords: Europe; income polarization; RIF-regression; relative distribution

JEL Classification: C14; D31; D63

Citation: Fabiani, Michele. 2023. Unraveling the Roots of Income Polarization in Europe: A Divided Continent. *Economies* 11: 217. https://doi.org/10.3390/economies11080217

Academic Editor: Franklin G. Mixon

Received: 16 May 2023
Revised: 5 July 2023
Accepted: 2 August 2023
Published: 16 August 2023

1. Introduction

Income polarization as a notion is comparatively understudied, particularly in Europe, compared to inequality and poverty, which have both gotten a lot of attention in the literature. Although both categories are sensitive to the middle of the distribution, income polarization and inequality are two distinct phenomena. While income inequality focuses on how far apart various members of a society are from the general mean, income polarization contrasts homogeneity within a group with the overall variability of a particular community (Castro 2003). Income polarization therefore resembles segregation more than income disparity (Esteban and Ray 1994).

The middle class is dwindling, which may be related to income polarization. Every civilization needs a prosperous middle class since it is linked to high income, rapid economic expansion, and social and political stability (Easterly 2001; Pressman 2007). High income polarization, on the other hand, suggests a divided society and may result in the creation of social conflict, discontent and tension (Esteban and Ray 1994, 1999; Gradín 2000; Zhang and Kanbur 2001). Although both income inequality and income polarization indicate shifts in the middle of the income distribution, income polarization is more likely to result in social unrest and political disorder.

In addition to causing societal discontent and conflict, income polarization can also have other negative effects. First of all, less social mobility results from a highly income-polarized society due to the fact that it may be challenging for the comparatively poor to advance up the income scale (Motiram and Sarma 2014). Polarization of income also has a negative impact on economic growth (Brzeziński 2013; Ezcurra 2009), affecting redistribution with possible negative effects, for example, on consumption. One explanation is that the social unrest and political unpredictability that underlie income polarization

could adversely affect market operations, labor relations, and the security of property rights (Keefer and Knack 2002). Moreover, income polarization is detrimental to health because it lowers the availability of some public goods and increases psychosocial stress due to social friction and conflict (Pérez and Ramos 2010).

This article contributes to the literature by observing the polarization trends in 12 European countries for the period from the early 2000s to the end of the second decade of the century. It provides a different view than previous works for two main reasons: it observes income distribution through polarization analysis instead of using traditional inequality measures, and it analyzes how household characteristics impact the distribution. To observe these trends, this paper uses the "relative distribution" method (Handcock and Morris 1998, 1999), a non-parametric approach. This methodology has not been used very frequently for the analysis of polarization and has been limited to single-country cases in Europe. This paper aims to provide a more complete picture of polarization analysis in the region, showing updated and harmonized data across countries. Furthermore, the relative distribution method is able to provide us with results that are easily usable and of immediate interpretation.

As a consequence, within the relative distribution framework, the paper applies a novel methodology to identify the covariates of distributional changes. The methodology used, i.e., the RIF-regression proposed by Firpo et al. (2009), is able to observe the impact of covariates on the distribution in detail, providing the possibility to assess economic policy interventions that should be taken to counteract the phenomenon of income polarization.

The paper is organized as follows: Section 2 provides a literature review on the issue of polarization in Europe; Section 3 discusses the data and provides summary statistics; Section 4 outlines the distinctive features of the relative distribution approach and presents the proposed RIF-regression approach; Section 5 details the main findings of the study; Section 6 provides summary conclusions.

2. Literature Review

Income polarization, defined as a divergence of income levels in a population, has been a topic of great interest in recent years, particularly outside Europe. Scholars have examined the matter in countries such as China (Araar 2008; Zhang and Kanbur 2001), India (Chakravarty and Majumder 2001; Motiram and Sarma 2014), Nigeria (Awoyemi and Araar 2009; Clementi et al. 2015), Sub-Saharan Africa (Clementi et al. 2019, 2021, 2022a) and Latin American countries (Deutsch et al. 2014; Gasparini et al. 2008), as well as in more developed countries like the United States and Canada (D'Ambrosio and Wolff 2001; Foster and Wolfson 1992, 2010). While there has been some research on income polarization in Europe, studies focused on this topic are relatively rare, and little attention has been paid to income polarization in new member states of Central and Eastern Europe (CEE NMS).

Most studies on income polarization in Europe use different approaches than the one proposed in this paper, in addition to analyzing the conditions of individual countries. Some of them use the polarization index DER, which identifies the formation of poles on the basis of within- and between-group effects. It is observed that in Italy and Spain (D'Ambrosio 2001; Gradín 2000), the specific characteristics of workers, such as education or area of residence, significantly influence the formation of these poles, thus leading to an increase in polarization. A similar approach has been used to analyze income polarization in Denmark. According to this study (Hussain 2009) the main factor that has led to an increase in this phenomenon is the between-group component, with a growing alienation of the people, who distance themselves more and more from each other, also bringing more inequality. For Poland (Brzeziński 2011), using the DER bi-polarization index, significant differences can be observed between the trend of inequality and that of polarization; the main component of the increase in polarization is found, in fact, in the growing identification within the groups themselves, leading to different trends between inequality and polarization.

Instead of analyzing the Italian situation, Poggi and Silber (2010) used the polarization index proposed by Deutsch et al. (2007); they showed that when taking the identity of

the individuals into account, a distinction can be made between a change over time in polarization and a change in polarization. The first one is the consequence of the so called "structural mobility", i.e., the change over time in the overall, between- and within-group inequality. The second one is the sole consequence of "exchange mobility", i.e., the changes over time in the ranks of individuals.

Analyzing the case of Germany, Gigliarano and Mosler (2009) used a multidimensional approach in the analysis of polarization, which no longer focuses on income polarization only, but also on education. The polarization index in this case will be a function of three components: inequality within groups, inequality between groups and the size of the groups themselves.

Atkinson and Brandolini (2013) used various measures of inequality to observe how changes in the distribution impact the middle class: a greater polarization on the tails of the distribution leads to a progressive emptying of the middle class, with both economic and social negative consequences. The results show that this effect is greater for southern European countries than for continental and northern European ones.

Finally, the relative distribution method has been used in some cases to observe income polarization in Russia and Italy (Nissanov and Pittau 2016; Massari et al. 2009b; Ricci and Scicchitano 2021). Studies show how during the analyzed periods, polarization grows in the country, even significantly. However, unlike our approach, these papers do not investigate what might be the driver that impacts on the changes in distribution leading to the observed growing polarization of income.

3. Data and Summary Statistics

In this paper, data are taken from the Luxembourg Income Study Database (LIS)[1]. LIS acquires datasets with income, wealth, employment and demographic data from many high- and middle-income countries, and harmonizes them to enable cross-national comparisons. Data used cover 12 European countries[2] available in the dataset and for which it is possible to have comparable surveys for two separate years.

The variable used in the first part for the distribution analysis is household-disposable income, net of income taxes and contributions[3].

To analyze the impact of social conditions on polarization trends in the second part of the paper, demographic, geographic, employment status and educational level of the head of household variables are used.

The period of analysis covers about two decades for all countries. Before turning to the analysis of polarization using the relative distribution method, it is interesting to look at some statistics regarding the trend of income inequality in the countries analyzed. Starting with income trends, is has been observed that the variable has shown growth, both for different income classes and on average among all the countries considered, except Spain and Italy. In fact, these countries show a stagnation of incomes, which remain practically stationary during the period taken into analysis.

It is also very interesting to observe the trend in the Gini index and the Foster-Wolfson polarization index, except in some specific cases, such as Germany and Luxembourg, which show a significant increase in inequality. For the remaining countries, the change is minimal or absent. In the United Kingdom, we even observe a decrease in these indexes.

4. Methodology

4.1. Polarization and Relative Distribution

In the analysis of income distribution, the topic of polarization has gained significance over the past 20 years (Foster and Wolfson 1992; Esteban and Ray 1994; Wolfson 1994, 1997), and it now appears that polarization is widely recognized as a separate concept from inequality.

Regardless of where a community is positioned along the income scale, a broad definition of income polarization (Esteban and Ray 1994) describes it as the "clustering" of a population around two or more poles of distribution. In a multi-group setting, the idea

of income polarization aims to quantify the degree of potential conflict present in a given distribution (see Esteban and Ray 1999, 2008, 2011). The concept is to think of society as a collection of groups, where members of one group have traits in common with one another (i.e., a sense of "identification"), but vary from members of other groups (i.e., a sense of "alienation") in terms of the same traits.

Therefore, political or social conflict is more likely the more homogeneous and separate the groups are; that is, when the within-group income distribution is more concentrated around its local mean, the between-group income distance is greater (see, inter alia, Gradín 2000; Milanovic 2000; D'Ambrosio 2001; Zhang and Kanbur 2001; Montalvo and Reynal-Querol 2002; Duclos et al. 2004; Lasso de la Vega and Urrutia 2006; Esteban et al. 2007; Gigliarano and Mosler 2009; Poggi and Silber 2010).

While summary measures of income polarization are frequently used in literature, a different (yet non-parametric) approach has emerged to measure the growth of the middle class and the level of household income polarization in a number of middle- and high-income nations. This approach is known as the "relative distribution" and combines the strengths of summary polarization indexes with the details of distributional change provided by the kernel density estimates.

The relative distribution method has been employed by Alderson et al. (2005), Massari et al. (2009a, 2009b), Alderson and Doran (2011), Borraz et al. (2013), Clementi and Schettino (2013, 2015), Clementi et al. (2015, 2017, 2019, 2021, 2022a), Molini and Paci (2015), Petrarca and Ricciuti (2016), and Nissanov and Pittau (2016).

More formally,[4] let Y_0 be the income variable for the reference population and Y the income variable for the comparison population. The relative distribution is defined as the ratio of the density of the comparison population to the density of the reference population, evaluated at the relative data r:

$$g(r) = \frac{f\left(F_0^{-1}(r)\right)}{f_0\left(F_0^{-1}(r)\right)} = \frac{f(y_r)}{f_0(y_r)},\ 0 \le r \le 1,\ y_r \ge 0, \tag{1}$$

where $f(\cdot)$ and $f_0(\cdot)$ denote the density functions of Y and Y_0, respectively, and $y_r = F_0^{-1}(r)$ is the quantile function of Y_0. When no changes occur between the two distributions, $g(r)$ has a uniform distribution. A value of $g(r)$ higher (lower) than 1 means that the share of households in the comparison population is higher (lower) than the corresponding share in the reference population at the rth quantile of the latter.

One of the major advantages of this method is the possibility to decompose the relative distribution into changes in location and shape. The decomposition can be written as:

$$\underbrace{\frac{f(y_r)}{f(y_r)}}_{Overall} = \underbrace{\frac{f_{0L}(y_r)}{f_0(y_r)}}_{Location} \times \underbrace{\frac{f(y_r)}{f_{0L}(y_r)}}_{Shape}. \tag{2}$$

$F_{0L}(y_r)$ is the median-adjusted density function:

$$f_{0L}(y_r) = f_0(y_r + \rho), \tag{3}$$

where the value ρ is the difference between the medians of the comparison and reference distributions; alternative indexes like the mean and/or multiplicative location shift can also be considered.

The relative distribution approach also includes a median relative polarization index, which is a measurement of the degree to which the comparison distribution is more polarized than the reference one:

$$MRP = \frac{4}{n}\left(\sum_{i=1}^{n}\left|r_i - \frac{1}{2}\right|\right) - 1. \tag{4}$$

The MRP index can be additively decomposed into contributions to the overall polarization made by the lower and upper halves of the median-adjusted relative distribution, enabling one to distinguish downgrading from upgrading. In terms of data, the lower relative polarization (LRP) index and the upper relative polarization (URP) index can be calculated as shown in Equations (5) and (6):

$$LRP = \frac{8}{n}\left[\sum_{i=1}^{n/2}\left(\frac{1}{2} - r_i\right)\right] - 1. \tag{5}$$

$$URP = \frac{8}{n}\left[\sum_{i=\frac{n}{2}+1}^{n}\left(r_i - \frac{1}{2}\right)\right] - 1. \tag{6}$$

with $MRP = \frac{1}{2}\ (LRP + URP)$. The *MRP*, *LRP* and *URP* indexes range from -1 to 1, and equal 0 when there is no change.

4.2. RIF-Regression Model

To analyze the drivers of income polarization in countries under consideration, the Recentered Influence Function (RIF) regression was used. The strength of the correlation between modest change in one covariate and change in a relative polarization index (such as MRP, LRP or URP) can be calculated using this method.

The influence function (Cowell and Victoria-Feser 1996) captures the effects of explanatory variables on the distributional statistic of interest, while also reflecting the influence of a single observation on a given distributional statistic, such as a particular quantile.

Firpo et al. (2009) proposed a simple modification in which the quantile is added back to the influence function, resulting in what the authors call the "re-centered influence function" (RIF):

$$RIF(y_i; q_\tau, F_Y) = q_\tau + IF(y_i; q_\tau, F_Y),$$

where q_τ is the τ-th quantile of the distribution of household incomes Y, and $IF(\cdot)$ is the influence function. With this result, Firpo et al. (2009) show that we can model the conditional expectation of the RIF as a simple linear function of explanatory variables.

In practice, following Firpo et al.'s (2009) procedure, one can first obtain an estimate of the RIF for each income *i* by using Equation (6) above; then, the following equation can be estimated using an ordinary least-squares method (OLS):

$$RIF(y_i; \hat{q}_\tau, F_Y) = \alpha_\tau + \sum_{k=1}^{K} \beta_{\tau,k} \cdot x_{\tau,i,k} + \varepsilon_{\tau,i}, \quad i = 1, 2, \ldots, N,$$

where α_τ is a constant, $x_{\tau,i,k}$ denotes a realization of the k-th explanatory variable, $\beta_{\tau,k}$ is the corresponding coefficient and $\varepsilon_{\tau,i}$ is the corresponding error term. The estimated model parameters $\hat{\beta}_{\tau,k}$, termed "unconditional quantile partial effect", can be interpreted as the effect of a small change in the distribution of X_k on the quintile q_τ—when the distribution of other covariates remains unchanged—or as a linear approximation of the effect of large changes of X_k on q_τ (e.g., Firpo et al. 2018).

Although Firpo et al. (2009) initially concentrated on the analysis of partial eects of explanatory variables on unconditional quantiles of the dependent variable, the underlying ideas of this methodology have been applied to other distributional statistics (for example, see Essama-Nssah and Lambert 2012; Rios-Avila 2020; Jann 2021)[5].

5. Results

5.1. Relative Distribution Results

There are two important aspects to consider. As noted in Table 1, the results obtained with this method are contrasted with the inequality results obtained with traditional measures (e.g., Gini index), where we do not observe a clear and common trend across countries. This underscores the importance of looking at the income distribution from dif-

ferent vantage points as well, to capture changes in the distribution, as relative distribution does. Relative distribution indexes show a homogenous pattern throughout the European countries surveyed. MRP, LRP and URP indexes are specified in Table 2, and graphs for each country, that allow an immediate and easy-to-read view of changes in distribution, are available in Appendix A. All are undergoing an accentuated polarization process, as evidenced by the positive and significant value of the MRP. The second aspect to consider is in which part of the distribution this concentration occurs. In 9 out of 11 countries (we exclude Italy from the analysis, where the index values are non-significant), the LRP index value is higher than the URP value, showing a more pronounced concentration in the lower tail than in the upper tail. This aspect is very significant, because it leads to a general worsening of the income distribution on the one hand, and on the other hand, to a progressive emptying of the middle class, which, as the results show, is sucked into the lower tail of the distribution.

Table 1. Summary statistics and inequality and polarization indexes.

Country	Year [a]	P10	P25	P50	Mean	P75	P90	Gini	Fgt0	FW
Austria	2000	15,453.1	21,222.0	28,213.5	30,736.8	36,628.5	48,616.9	25.4	13.7	20.1
	2019	17,059.0	25,105.6	34,097.5	37,783.3	46,174.6	61,489.0	27.4	15.4	21.7
Belgium	2000	14,104.1	18,805.4	27,002.7	30,313.4	35,868.5	47,013.6	28.8	16.2	22.0
	2017	15,335.6	21,104.6	30,880.9	32,847.5	40,884.2	51,391.1	26.0	18.4	22.1
Denmark	2000	15,953.2	20,758.6	28,254.5	29,686.8	35,772.6	43,914.9	22.5	13.1	18.2
	2016	17,606.8	22,771.4	31,255.8	34,168.6	41,277.4	52,309.4	25.5	12.8	20.5
Finland	2000	12,605.8	16,324.7	22,272.2	24,495.9	28,995.4	36,976.2	25.3	12.7	20.0
	2016	15,827.0	20,916.9	28,139.0	31,381.4	37,410.7	48,046.0	25.8	12.6	20.4
France	2000	12,744.5	17,415.9	24,065.7	27,858.2	33,224.6	46,376.4	29.4	14.9	23.6
	2018	13,815.4	19,275.9	26,970.8	31,095.5	36,605.6	50,458.9	30.2	16.0	23.0
Germany	2000	15,184.2	20,825.9	27,538.6	30,606.1	36,915.0	48,383.3	25.9	12.5	20.7
	2019	15,291.0	22,182.4	31,483.6	35,222.0	42,200.6	56,566.9	29.3	17.2	22.8
Ireland	2000	9322.0	14,279.8	22,236.7	25,007.8	31,176.6	41,731.6	31.3	22.5	26.4
	2019	16,008.1	21,494.2	30,308.7	34,853.2	42,441.0	55,680.9	28.7	15.5	23.8
Italy	2000	8945.6	13,425.4	20,400.7	23,793.6	29,665.8	40,105.7	33.4	20.1	28.2
	2016	8206.8	12,741.7	19,503.7	22,359.3	28,518.8	39,064.8	33.9	21.1	29.1
Luxembourg	2000	21,288.6	27,628.0	37,282.0	42,403.1	51,239.9	69,413.8	26.2	12.3	22.8
	2019	21,967.5	30,451.9	43,198.6	49,813.8	61,648.5	82,427.8	29.6	16.4	25.5
Netherland	1999	15,596.0	20,104.8	26,656.0	28,670.3	34,764.7	43,537.3	23.1	11.1	19.0
	2018	16,802.5	22,378.2	30,833.0	34,284.5	41,330.4	53,688.4	27.0	13.8	21.6
Spain	2000	9857.0	14,486.5	22,246.8	26,265.5	32,556.0	46,205.8	33.7	20.8	29.2
	2016	8852.7	14,606.6	23,047.6	26,407.8	34,173.7	46,516.1	34.1	22.6	29.6
United Kingdom	2000	10,543.7	14,622.7	22,152.3	27,690.4	32,969.3	47,166.8	35.7	20.3	29.5
	2020	14,228.5	19,148.3	27,222.5	31,741.7	38,876.6	54,271.9	30.5	15.5	25.8

Source: Authors' calculations based on LIS data. [a] The time frame used for different countries is determined by data availability.

Table 2. Polarization indexes by country.

Country	Index [a]	Value	LB [b]	UB [c]	*p*-Value [d]
Austria	MRP	0.150	0.113	0.187	0.000
	LRP	0.161	0.099	0.222	0.000
	URP	0.139	0.091	0.188	0.000
Belgium	MRP	0.102	0.065	0.139	0.000
	LRP	0.148	0.084	0.212	0.000
	URP	0.056	0.003	0.109	0.036

Table 2. *Cont.*

Country	Index [a]	Value	LB [b]	UB [c]	*p*-Value [d]
Denmark	MRP	0.127	0.120	0.134	0.000
	LRP	0.101	0.089	0.113	0.000
	URP	0.153	0.143	0.163	0.000
Finland	MRP	0.155	0.134	0.177	0.000
	LRP	0.157	0.117	0.197	0.000
	URP	0.153	0.126	0.181	0.000
France	MRP	0.065	0.055	0.075	0.000
	LRP	0.105	0.089	0.121	0.000
	URP	0.026	0.014	0.038	0.000
Germany	MRP	0.148	0.127	0.168	0.000
	LRP	0.213	0.178	0.247	0.000
	URP	0.082	0.054	0.111	0.000
Ireland	MRP	0.120	0.070	0.170	0.000
	LRP	0.090	0.001	0.180	0.047
	URP	0.150	0.088	0.211	0.000
Italy	MRP	−0.005	−0.038	0.028	0.772
	LRP	0.007	−0.054	0.068	0.823
	URP	−0.017	−0.058	0.024	0.416
Luxembourg	MRP	0.178	0.130	0.227	0.000
	LRP	0.234	0.146	0.321	0.000
	URP	0.122	0.065	0.179	0.000
Netherland	MRP	0.167	0.141	0.193	0.000
	LRP	0.188	0.143	0.232	0.000
	URP	0.146	0.111	0.181	0.000
Spain	MRP	0.045	0.016	0.075	0.002
	LRP	0.081	0.030	0.132	0.002
	URP	0.010	−0.025	0.046	0.569
United Kingdom	MRP	0.058	0.030	0.085	0.000
	LRP	0.083	0.035	0.132	0.001
	URP	0.032	0.002	0.061	0.031

Source: Authors' calculations based on LIS data. Notes: [a] MRP = median relative polarization index; LRP = lower relative polarization index; URP = upper relative polarization index. [b] Lower bound of the 95% confidence interval. [c] Upper bound of the 95% confidence interval. [d] Refers to the null hypothesis of no change with respect to the reference distribution, i.e., that the index equals to 0.

5.2. RIF-Regression Results

In this section are presented the results of RIF-regressions for the three polarization indexes and different independent variables. The independent variables are divided into different categories, such as the sector of employment, education level, country of birth and area of residence. The choice of variables is related to their availability within the surveys and following what has been carried out in previous analyses of income polarization (e.g., Ezcurra 2009). Each category has a base group, against which the other groups are compared. The coefficients and standard errors for each independent variable are presented in the tables. Asterisks next to a coefficient indicate the level of statistical significance of that variable: *** denotes significance at the 1% level, ** at the 5% level and * at the 10% level.

Table 3 presents the results of a RIF-regression analysis, with the MRP index as the dependent variable and several independent variables. The independent variables are divided into different categories, and each category has a base group against which the other groups are compared. The results indicate that the sector of employment, education level, country of birth, area of residence and age have significant effects on the median relative polarization in some countries. For example, being employed in the agricultural sector has a positive effect on the median relative polarization in Austria and Belgium,

while the industry sector has a negative effect in some countries. Education level has also a significant effect, with those having a high level of education showing a positive effect on the median relative polarization in all countries. The country of birth and area of residence also have significant effects, with being born outside the country and living in rural areas having negative effects on the median relative polarization.

Table 4 shows the results of a regression analysis with the dependent variable being the lower relative polarization index of income, and several independent variables are used to explain the variation in the dependent variable across different sectors and countries. The independent variables are grouped into four categories: sector, education, country of birth and area. The coefficients of these binary variables indicate the effect of being in that subcategory on the dependent variable compared to the base category.

Table 5 displays the results of a regression analysis where the dependent variable is the upper relative polarization index of income, and various independent variables are examined across different countries. The results show that workers in the agricultural sector have a positive and significant effect on the upper relative polarization index of income compared to the not-employed group, while workers in the industry and service sectors have a negative and significant effect. Education level also has a significant effect, with workers having a high education level showing a positive and significant effect on the index compared to those with a low education level. Being born outside the country has a negative and significant effect on the index, while living in rural areas has a negative and significant effect on the index.

Table 3. Rif-regression results, MRP index.

	Austria	Belgium	Denmark	Finland	France	Germany	Ireland	Italy	Luxembourg	Netherland	Spain	U.K.
Sector												
Not employed	(base)	(base)	(base)	(base)	(base)	(base)	(base)	(base)	(base)	(base)	(base)	(base)
Agriculture	0.090 ** (0.040)	0.095 ** (0.039)	0.926 *** (0.247)	0.171 *** (0.050)	0.029 (0.173)	−0.217 (0.214)	0.047 * (0.025)	0.131 *** (0.043)	0.021 (0.053)	0.195 *** (0.061)	0.078 * (0.041)	0.088 (0.313)
Industry	0.004 (0.018)	−0.002 (0.014)	0.656 *** (0.084)	0.066 ** (0.029)	−0.327 *** (0.072)	−0.062 (0.043)	0.005 (0.013)	0.000 (0.033)	0.008 (0.016)	−0.096 *** (0.028)	−0.034 * (0.020)	0.360 *** (0.059)
Services	0.002 (0.012)	−0.015 (0.010)	0.672 *** (0.050)	0.033 (0.022)	−0.306 *** (0.057)	−0.021 (0.029)	0.008 (0.009)	−0.010 (0.024)	0.025 ** (0.012)	−0.049 *** (0.016)	−0.021 (0.016)	0.218 *** (0.042)
Education												
Low	(base)	(base)	(base)	(base)	(base)	(base)	(base)	(base)	(base)	(base)	(base)	(base)
Medium	−0.001 (0.013)	−0.019 * (0.009)	−0.242 *** (0.050)	−0.057 ** (0.024)	0.230 *** (0.052)	−0.150 *** (0.035)	0.039 *** (0.010)	0.098 *** (0.019)	0.015 (0.010)	−0.020 (0.016)	0.099 *** (0.016)	0.064 (0.047)
High	0.080 *** (0.017)	0.053 *** (0.009)	1.127 *** (0.057)	0.200 *** (0.025)	2.229 *** (0.062)	0.209 *** (0.039)	0.106 *** (0.009)	0.358 *** (0.034)	0.096 *** (0.011)	0.176 *** (0.017)	0.309 *** (0.015)	0.469 *** (0.043)
Country of birth												
Born in the country	(base)	(base)			(base)	(base)	(base)	(base)	(base)	(base)	(base)	(base)
Born outside the country	−0.014 (0.013)	0.017 *** (0.010)			0.335 *** (0.061)	0.039 (0.033)	−0.055 *** (0.009)	−0.173 *** (0.040)	−0.016 * (0.009)	0.074 *** (0.023)	−0.020 (0.020)	−0.041 (0.0549)
Area												
Cities	(base)	(base)	(base)	(base)	(base)		(base)		(base)		(base)	
Towns and suburbs	−0.008 (0.012)	−0.021 *** (0.008)	−0.650 *** (0.052)	−0.108 *** (0.021)	−0.461 *** (0.076)		−0.024 ** (0.010)		−0.022 (0.013)		−0.057 *** (0.015)	
Rural areas	−0.031 ** (0.012)	−0.034 *** (0.011)	−0.609 *** (0.081)	−0.117 *** (0.022)	−0.570 *** (0.087)		−0.041 *** (0.008)		−0.017 (0.013)		−0.102 *** (0.013)	
Age	0.001 *** (0.000)	−0.000 (0.000)	0.009 *** (0.001)	0.001 (0.001)	0.010 *** (0.001)	−0.003 *** (0.001)	0.001 *** (0.000)	0.000 (0.000)	0.001 ** (0.000)	0.000 (0.000)	0.001 * (0.000)	0.000 (0.001)

Table 3. *Cont.*

	Austria	Belgium	Denmark	Finland	France	Germany	Ireland	Italy	Luxembourg	Netherland	Spain	U.K.
Sex												
Male	(base)	(base)	(base)	(base)	(base)	(base)	(base)	(base)	(base)	(base)	(base)	(base)
Female	0.016 (0.010)	0.002 (0.007)	0.219 *** (0.048)	−0.074 *** (0.018)	−0.329 *** (0.049)	−0.066 *** (0.025)	−0.003 (0.007)	−0.053 *** (0.016)	−0.003 (0.009)	−0.040 *** (0.014)	−0.034 *** (0.012)	0.155 *** (0.036)
N° Household members	−0.001 (0.004)	−0.000 (0.003)	−0.125 *** (0.020)	−0.053 *** (0.007)	−0.016 (0.019)	−0.050 *** (0.009)	−0.003 (0.002)	0.019 ** (0.008)	−0.007 ** (0.003)	−0.014 *** (0.005)	0.014 *** (0.004)	−0.067 *** (0.016)

Notes: Robust standard errors in brackets; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 4. Rif-regression results, LRP index.

	Austria	Belgium	Denmark	Finland	France	Germany	Ireland	Italy	Luxembourg	Netherland	Spain	U.K.
Sector												
Not employed	(base)	(base)	(base)	(base)	(base)	(base)	(base)	(base)	(base)	(base)	(base)	(base)
Agriculture	0.118 ** (0.059)	0.123 (0.077)	1.906 *** (0.427)	0.267 *** (0.088)	−0.633 ** (0.298)	−0.477 (0.408)	0.093 ** (0.042)	0.167 ** (0.076)	−0.007 (0.095)	0.236 ** (0.103)	0.081 (0.076)	−0.193 (0.489)
Industry	0.004 (0.030)	0.022 (0.024)	2.232 *** (0.151)	0.152 *** (0.051)	−0.634 *** (0.122)	−0.080 (0.075)	0.026 (0.024)	0.061 (0.061)	0.020 (0.032)	−0.095 * (0.051)	0.019 (0.037)	0.744 *** (0.105)
Services	−0.010 (0.020)	0.012 (0.018)	2.139 *** (0.090)	0.046 (0.040)	−0.853 *** (0.096)	0.008 (0.047)	0.025 (0.016)	0.055 (0.046)	0.056 ** (0.023)	−0.068 ** (0.028)	−0.011 (0.029)	0.311 *** (0.079)
Education												
Low	(base)	(base)	(base)	(base)	(base)	(base)	(base)	(base)	(base)	(base)	(base)	(base)
Medium	−0.004 (0.024)	−0.001 (0.017)	0.025 (0.096)	−0.067 (0.045)	0.479 *** (0.097)	−0.196 *** (0.064)	0.066 *** (0.019)	0.152 *** (0.038)	0.032 (0.020)	−0.055 * (0.029)	0.166 *** (0.031)	0.103 (0.088)
High	0.063 ** (0.030)	0.088 *** (0.017)	2.069 *** (0.104)	0.245 *** (0.045)	2.401 *** (0.104)	0.225 *** (0.068)	0.145 *** (0.018)	0.374 *** (0.054)	0.158 *** (0.020)	0.180 *** (0.029)	0.390 *** (0.027)	0.617 *** (0.079)
Country of birth												
Born in the country	(base)	(base)			(base)	(base)	(base)	(base)	(base)	(base)	(base)	(base)

Table 4. *Cont.*

	Austria	Belgium	Denmark	Finland	France	Germany	Ireland	Italy	Luxembourg	Netherland	Spain	U.K.
Born outside the country	−0.019 (0.023)	−0.025 (0.019)			0.354 *** (0.105)	0.018 (0.059)	−0.092 *** (0.017)	−0.344 *** (0.075)	−0.056 *** (0.016)	0.066 (0.040)	−0.097 ** (0.039)	−0.132 (0.098)
Area												
Cities	(base)	(base)	(base)	(base)	(base)		(base)		(base)		(base)	
Towns and suburbs	−0.010 (0.019)	−0.022 (0.015)	−0.464 *** (0.089)	−0.083 ** (0.038)	−0.638 *** (0.141)		−0.035 ** (0.017)		−0.028 (0.023)		−0.059 ** (0.028)	
Rural areas	−0.057 *** (0.019)	−0.037 * (0.020)	−0.192 *** (0.146)	−0.079 ** (0.040)	−0.780 *** (0.156)		−0.047 *** (0.015)		−0.013 (0.024)		−0.123 *** (0.026)	
Age	0.000 (0.000)	−0.001 ** (0.000)	0.000 (0.002)	−0.001 (0.001)	−0.004 * (0.002)	−0.005 *** (0.001)	0.001 ** (0.000)	0.000 (0.001)	0.001 *** (0.001)	−0.003 *** (0.000)	0.001 (0.001)	−0.001 (0.002)
Sex												
Male	(base)	(base)	(base)	(base)	(base)	(base)	(base)	(base)	(base)	(base)	(base)	(base)
Female	0.013 (0.016)	−0.014 (0.013)	0.079 (0.088)	−0.077 ** (0.032)	−0.496 *** (0.086)	−0.047 (0.045)	0.001 (0.013)	−0.040 (0.031)	0.004 (0.017)	−0.089 *** (0.025)	−0.044 * (0.022)	−0.226 *** (0.066)
N° Household members	−0.002 (0.006)	0.006 (0.005)	0.114 *** (0.034)	−0.079 *** (0.012)	−0.045 (0.034)	−0.045 *** (0.016)	−0.005 (0.004)	0.006 (0.014)	−0.011 (0.007)	−0.013 (0.009)	0.029 *** (0.008)	−0.108 *** (0.029)

Notes: Robust standard errors in brackets; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 5. Rif-regression results, URP index.

	Austria	Belgium	Denmark	Finland	France	Germany	Ireland	Italy	Luxembourg	Netherland	Spain	U.K.
Sector												
Not employed	(base)	(base)	(base)	(base)	(base)	(base)	(base)	(base)	(base)	(base)	(base)	(base)
Agriculture	0.062 (0.054)	0.067 * (0.040)	−0.054 (0.353)	0.075 (0.068)	0.692 *** (0.191)	0.041 (0.163)	0.000 (0.033)	0.094 ** (0.041)	0.049 (0.046)	0.154 * (0.084)	0.075 * (0.043)	0.371 (0.232)
Industry	0.005 (0.026)	−0.027 (0.019)	−0.920 *** (0.119)	−0.020 (0.035)	−0.019 (0.089)	−0.044 (0.058)	−0.015 (0.018)	−0.062 (0.040)	−0.002 (0.014)	−0.097 ** (0.043)	−0.088 *** (0.027)	−0.023 (0.072)

Table 5. *Cont.*

	Austria	Belgium	Denmark	Finland	France	Germany	Ireland	Italy	Luxembourg	Netherland	Spain	U.K.
Services	0.014 (0.015)	−0.043 *** (0.012)	−0.794 *** (0.067)	0.019 (0.025)	0.240 *** (0.064)	−0.051 (0.044)	−0.008 (0.011)	−0.077 *** (0.024)	−0.004 (0.012)	−0.030 (0.021)	−0.031 (0.019)	0.125 *** (0.042)
Education												
Low	(base)	(base)	(base)	(base)	(base)	(base)	(base)	(base)	(base)	(base)	(base)	(base)
Medium	0.002 (0.015)	−0.037 *** (0.012)	−0.510 *** (0.067)	−0.048 * (0.027)	−0.019 (0.058)	−0.104 ** (0.052)	0.012 (0.012)	0.044 ** (0.022)	−0.001 (0.010)	0.015 (0.021)	0.032 (0.020)	0.025 (0.052)
High	0.097 *** (0.021)	−0.017 (0.012)	0.186 ** (0.078)	0.155 *** (0.029)	2.057 *** (0.073)	0.193 *** (0.058)	0.066 *** (0.011)	0.342 *** (0.036)	0.034 *** (0.012)	0.171 *** (0.023)	0.229 *** (0.019)	0.322 *** (0.048)
Country of birth												
Born in the country	(base)	(base)			(base)	(base)	(base)	(base)	(base)	(base)	(base)	(base)
Born outside the country	−0.010 (0.016)	0.061 *** (0.012)			0.317 *** (0.067)	0.060 (0.045)	−0.017 * (0.010)	−0.001 (0.031)	0.022 * (0.011)	0.082 *** (0.026)	0.056 ** (0.023)	0.050 (0.056)
Area												
Cities	(base)	(base)	(base)	(base)	(base)		(base)		(base)		(base)	
Towns and suburbs	−0.007 (0.017)	−0.020 * (0.011)	−0.836 *** (0.072)	−0.134 *** (0.025)	−0.284 *** (0.085)		−0.014 (0.012)		−0.016 (0.015)		−0.055 *** (0.019)	
Rural areas	−0.005 (0.016)	−0.031 ** (0.015)	−1.025 *** (0.115)	−0.156 *** (0.026)	−0.360 *** (0.107)		−0.035 *** (0.010)		−0.021 (0.015)		−0.082 *** (0.018)	
Age	0.001 *** (0.001)	0.001 * (0.001)	0.019 *** (0.001)	0.002 *** (0.001)	0.025 *** (0.001)	−0.001 (0.001)	0.001 * (0.000)	0.001 ** (0.001)	0.000 (0.000)	0.002 *** (0.001)	0.001 ** (0.000)	0.001 (0.001)
Sex												
Male	(base)	(base)	(base)	(base)	(base)	(base)	(base)	(base)	(base)	(base)	(base)	(base)
Female	0.018 (0.012)	0.019 * (0.011)	0.359 *** (0.068)	−0.071 *** (0.021)	−0.162 *** (0.050)	−0.086 ** (0.035)	−0.006 (0.009)	−0.066 *** (0.020)	−0.010 (0.011)	0.008 (0.018)	−0.025 (0.016)	−0.084 ** (0.039)
N° Household members	−0.000 (0.005)	−0.008 * (0.004)	−0.364 *** (0.028)	−0.026 *** (0.009)	0.012 (0.022)	−0.055 *** (0.014)	−0.001 (0.003)	0.033 *** (0.008)	−0.003 (0.003)	−0.016 ** (0.006)	−0.001 (0.006)	−0.027 (0.016)

Notes: Robust standard errors in brackets; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

6. Discussion

The research presented is intended to provide new insights from previous analyses of inequality and polarization in Europe. Starting precisely with the distinction between the analysis of income inequality and polarization, the paper adds to the scant literature present for the latter issue (D'Ambrosio 2001; Gradín 2000; Hussain 2009; Brzeziński 2011; Poggi and Silber 2010). Further, it observes polarization using a method (i.e., relative distribution) hitherto sparsely employed for European countries, with few exceptions (Petrarca and Ricciuti 2016; Nissanov and Pittau 2016), which allows us to observe how incomes are distanced from each other from an absolute rather than relative point of view (Clementi et al. 2022b), as shown by other indexes proposed in previous research.

Turning to the second part of the analysis, the comparability of the country surveys used and the innovative RIF-regression method (Firpo et al. 2009) allow us to propose a general European picture of the causes of the polarization seen earlier. Such a comparison has not been performed in such detail before, opening up possible discussions in this regard. For one, the method used allows us to observe the impact of the variables considered on income distribution in a more timely manner. Second, having carried out the same analysis for different European countries, the results open up a policy discussion, leading to an analysis of what might be common policies to counteract this growing income distancing observed previously.

7. Conclusions

Income polarization is a concept that is gaining increasing importance in the analysis of income distribution in Europe. Although often confused with income inequality, income polarization is a distinct phenomenon that focuses on the homogeneity within a group rather than the differences between groups. Reduction in the middle class is a significant factor that may contribute to income polarization. High levels of income polarization can lead to social unrest, political instability and economic downturns. This article has used the relative distribution method and econometric decomposition to examine the trends in income polarization in 12 European countries over the past two decades. The analysis has identified the main drivers of polarization and their impact on observable and unobservable characteristics. Data used in the analysis were obtained from the Luxembourg Income Study Database (LIS), and cover income, wealth, employment and demographic details. The paper has contributed to the literature by providing a granular analysis of distributional changes that an analysis based on standard inequality decompositions would not allow. The results of the analysis show that polarization has increased over the past two decades in most countries examined, and the main drivers of this trend have been changes in the labor market and the level of education. The implications of these findings are significant, as income polarization has far-reaching effects on society, economy and individual well-being. A highly polarized society can result in the creation of social conflict, discontent and tension. It can also lead to lack of social mobility, lower economic growth, and can have a negative impact on health outcomes. Policies aimed at reducing income polarization are therefore necessary to promote social and political stability, economic growth and individual well-being. The results of this paper confirm what has been seen in previous studies regarding polarization in specific European countries: the phenomenon is primarily driven by the different conditions of households in the labor market (Gigliarano and Mosler 2009; Brzeziński 2013). This can thus be seen as the area where most action needs to be taken to ensure better income distribution. The most important difference between this paper and previous ones concerns the methodology used; relative distribution allows us to observe differences in absolute terms (Clementi et al. 2022a), thus ensuring a different point of view and one that may open up new discussions. It is very important to find common aspects between countries which, although on the same continent, are very different from a social, economic and dimensional point of view. The results clearly show the need to intervene first and above all, in the labor market, guaranteeing greater participation and support to people in difficulty or in a state of unemployment. The second

fundamental point of intervention is the education system: a higher level of education guarantees the possibility of accessing a higher level of income. On one hand, this implies guaranteeing new generations the possibility of improving and increasing their knowledge. On the other hand, it allows those already in the labor market and with lower levels of education the possibility to increase their skills through specific trainings focused on the emerging needs and trends of the labor market. Further research is needed to explore the complex relationships between income polarization, inequality and poverty, and to effectively identify policy measures to address these issues.

Funding: This research received no external funding.

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: The Luxembourg Income Study Database (https://www.lisdatacenter.org/, accessed on 10 March 2023) provides remote access to the microdata through a web-based Job Submission Interface (LISSY). Users have to register to the platform and submit through the LISSY interface their statistical programs written in R, SAS, SPSS or Stata.

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

Appendix A

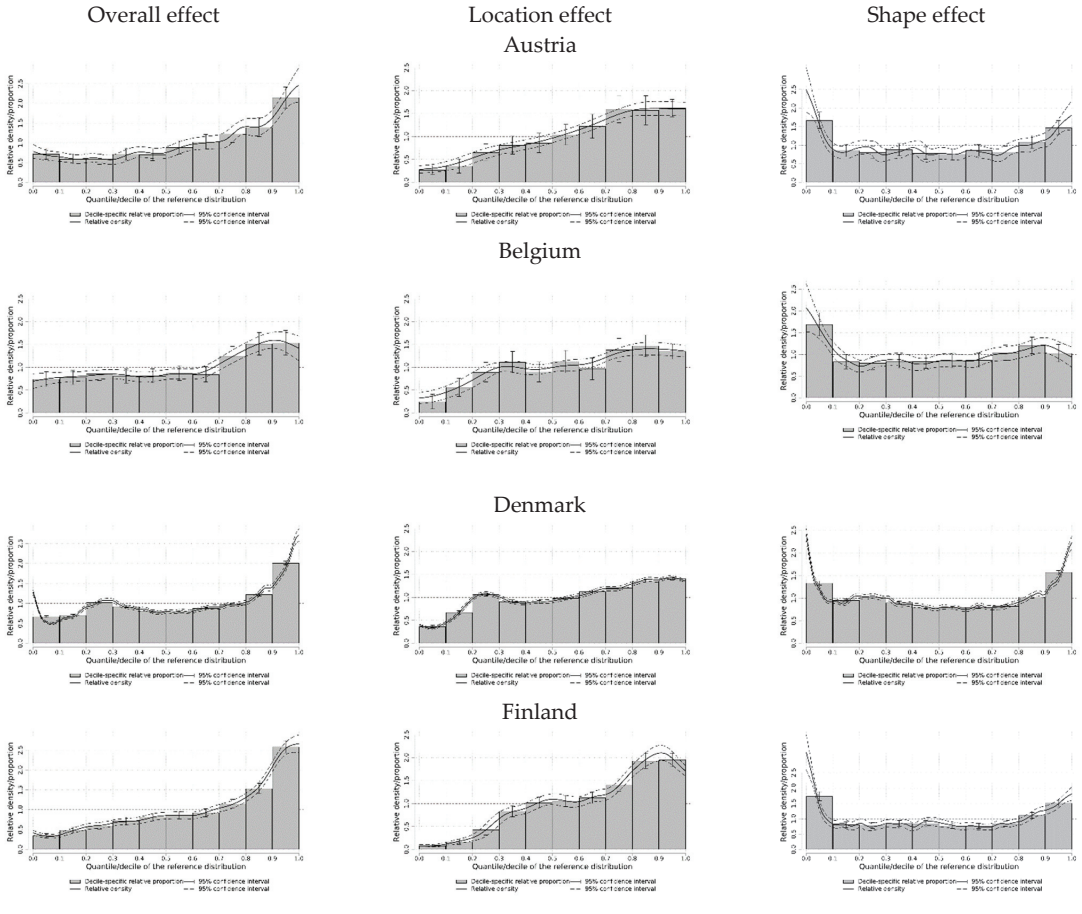


Figure A1. *Cont*.

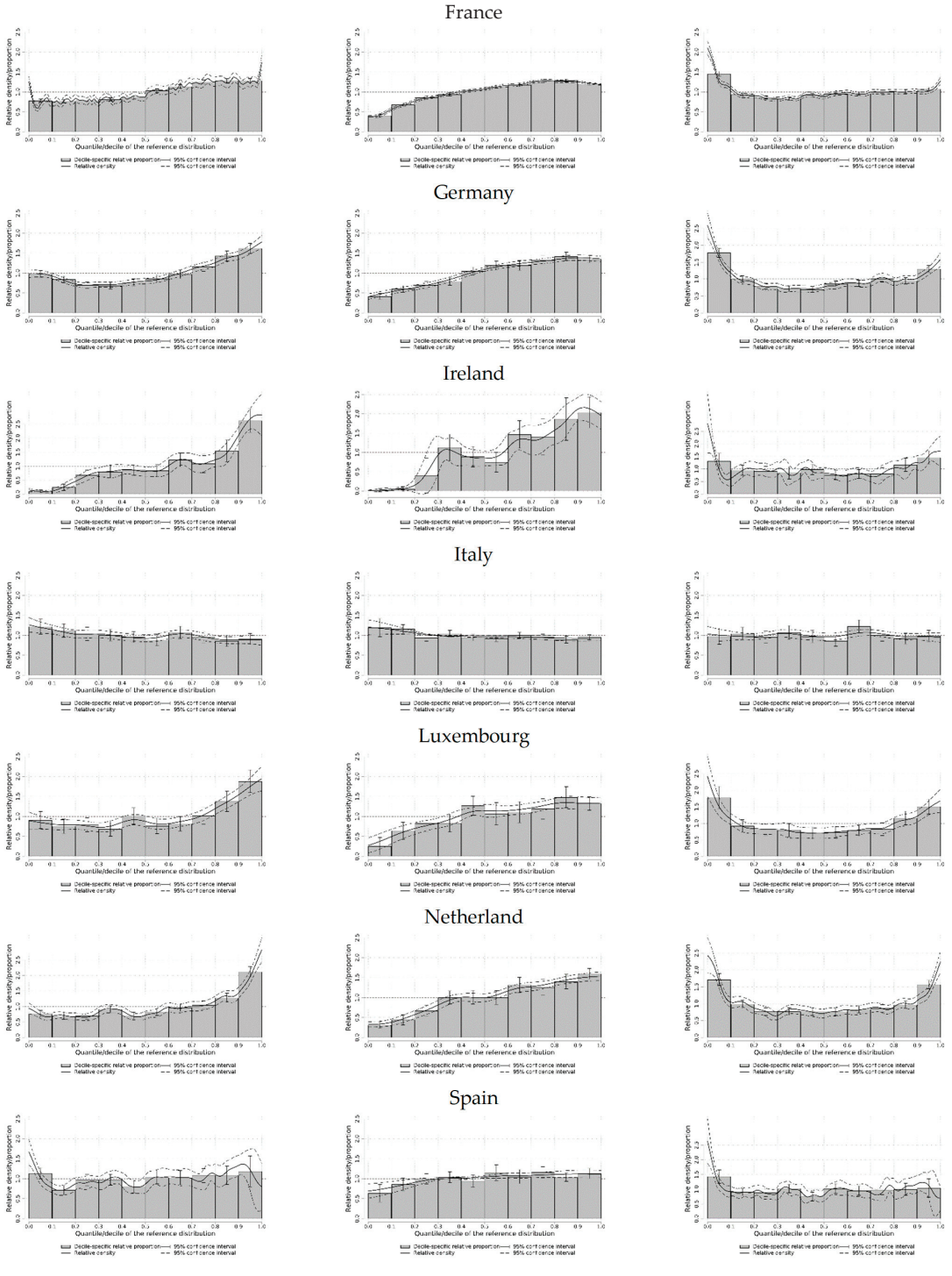


Figure A1. *Cont*.

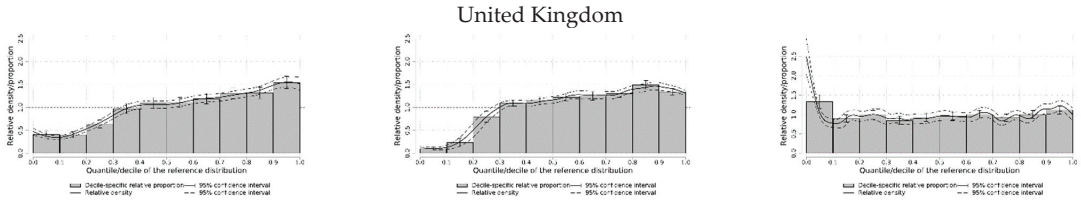


Figure A1. Source: Authors' calculations based on LIS data.

Notes

1 Luxembourg Income Study (LIS) Database, http://www.lisdatacenter.org (Accessed on 10 March 2023) (multiple countries; December 2022–January 2023). Luxembourg: LIS.

2 Austria, Belgium, Denmark, Finland, France, Germany, Ireland, Italy, Luxembourg, Netherland, Spain and United Kingdom.

3 "Disposable household income" is usually the preferred measure for income distribution analysis, as it is the income available to households to support their consumption expenditure and savings during the reference period (Canberra Group 2011). According to the LIS documentation (https://www.lisdatacenter.org/data-access/key-figures/methods/disposable/) Accessed on 10 March 2023, this measure includes income received from work, wealth and direct government benefits, such as retirement or unemployment benefits. The measure then subtracts direct taxes paid, such as income taxes.

4 Here, we limit ourselves to illustrating the basic concepts behind the use of the relative distribution method. Interested readers are referred to Handcock and Morris (1998, 1999) for a more detailed explication.

5 For a more specific observation of the use of RIF-regression applied to polarization indexes, see (Jann 2021).

References

Alderson, Arthur S., and Kevin Doran. 2011. Global inequality, within-nation inequality, and the changing distribution of income in seven transitional and middle-income societies. In *Inequality beyond Globalization: Economic Changes, Social Transformations, and the Dynamics of Inequality*. Münster: LIT Verlag, pp. 183–200.

Alderson, Arthur S., Jason Beckfield, and François Nielsen. 2005. Exactly How Has Income Inequality Changed? Patterns of Distributional Change in Core Societies. *International Journal of Comparative Sociology* 46: 405–23. [CrossRef]

Araar, Abdelkrim. 2008. On the Decomposition of Polarization Indices: Illustrations with Chinese and Nigerian Household Surveys. CIRPEE Working Paper No. 08-06. Montréal: CIRPEE.

Atkinson, Anthony B., and Andrea Brandolini. 2013. On the identification of the middle class. In *Income Inequality: Economic Disparities and the Middle Class in Affluent Countries*. Stanford: Stanford University Press, pp. 77–100.

Awoyemi, T. T., and A. Araar. 2009. Explaining polarization and its dimensions in Nigeria: A DER decomposition approach. Paper presented at 14th Annual Conference on 'Econometric Modelling for Africa', Abuja, Nigeria, July 8–10; pp. 8–10.

Borraz, Fernando, Nicolás González, and Maximo Rossi. 2013. Polarization and the middle class in Uruguay. *Latin American Journal of Economics* 50: 289–326. [CrossRef]

Brzeziński, Michał. 2011. Statistical inference on changes in income polarization in Poland. *Przegląd Statystyczny* 58: 102–13.

Brzeziński, Michal. 2013. *Income Polarization and Economic Growth*. National Bank of Poland Working Paper 147. Warsaw: National Bank of Poland.

Canberra Group. 2011. *Handbook on Household Income Statistics*. Geneva: United Nations.

Castro, José Villaverde. 2003. Regional Convergence, Polarisation and Mobility in the European Union, 1980–1996. *Journal of European Integration* 25: 73–86. [CrossRef]

Chakravarty, Satya R., and Amita Majumder. 2001. Inequality, polarisation and welfare: Theory and applications. *Australian Economic Papers* 40: 1–13. [CrossRef]

Clementi, Fabio, and Francesco Schettino. 2013. Income polarization in Brazil, 2001–2011: A distributional analysis using PNAD data. *Economics Bulletin* 33: 16. [CrossRef]

Clementi, Fabio, and Francesco Schettino. 2015. Declining inequality in Brazil in the 2000s: What is hidden behind? *Journal of International Development* 27: 929–52. [CrossRef]

Clementi, Fabio, Andrew L. Dabalen, Vasco Molini, and Francesco Schettino. 2015. When the Centre Cannot Hold: Patterns of Polarization in Nigeria. *Review of Income and Wealth* 63: 608–32. [CrossRef]

Clementi, Fabio, Michele Fabiani, and Vasco Molini. 2019. The devil is in the detail: Growth, inequality and poverty reduction in Africa in the last two decades. *Journal of African Economies* 28: 408–34. [CrossRef]

Clementi, Fabio, Michele Fabiani, and Vasco Molini. 2021. How polarized is sub-Saharan Africa? A look at the regional distribution of consumption expenditure in the 2000s. *Oxford Economic Papers* 73: 796–819. [CrossRef]

Clementi, Fabio, Michele Fabiani, Vasco Molini, and Francesco Schettino. 2022a. Response to 'Polarization and Poverty Reduction in Africa: The Devil Is in the Choice of Equivalence Relation'. *Journal of African Economies* 31: 176–79. [CrossRef]

Clementi, Fabio, Michele Fabiani, Vasco Molini, and Rocco Zizzamia. 2022b. Are we really painting the devil on the walls? Polarization and its drivers in Sub-Saharan Africa in the past two decades. *Journal of African Economies* 31: 124–46. [CrossRef]
Clementi, Fabio, Vasco Molini, and Francesco Schettino. 2017. All that Glitters is Not Gold: Polarization Amid Poverty Reduction in Ghana. *World Development* 102: 275–91. [CrossRef]
Cowell, Frank A., and Maria-Pia Victoria-Feser. 1996. Robustness properties of inequality measures. *Econometrica: Journal of the Econometric Society* 64: 77–101. [CrossRef]
de la Vega, Ma Casilda Lasso, and Ana Marta Urrutia. 2006. An alternative formulation of the Esteban-Gradín-Ray extended measure of polarization. *Journal of Income Distribution* 15: 42–54. [CrossRef]
D'Ambrosio, Conchita. 2001. Household Characteristics and the Distribution of Income in Italy: An Application of Social Distance Measures. *Review of Income and Wealth* 47: 43–64. [CrossRef]
D'Ambrosio, Conchita, and Edward N. Wolff. 2001. *Is Wealth Becoming More Polarized in the United States?* Jerome Levy Economics Institute of Bard College Working Paper 330. Annandale-On-Hudson: Jerome Levy Economics Institute of Bard College.
Deutsch, Joseph, Jacques Silber, and Gaston Yalonetzky. 2014. On Bi-Polarization and The Middle Class in L atin A merica: A Look A t the First Decade of the Twenty-First Century. *Review of Income and Wealth* 60: S332–S352. [CrossRef]
Deutsch, Joseph, Meital Hanoka, and Jacques Silber. 2007. On the Link Between the Concepts of Kurtosisand Bipolarization. *Economics Bulletin* 4: 1–6.
Duclos, Jean-Yves, Joan Esteban, and Debraj Ray. 2004. Polarization: Concepts, measurement, estimation. *Econometrica* 72: 1737–72. [CrossRef]
Easterly, William. 2001. The Middle Class Consensus and Economic Development. *Journal of Economic Growth* 6: 317–35. [CrossRef]
Essama-Nssah, Boniface, and Peter J. Lambert. 2012. Chapter 6 Influence Functions for Policy Impact Analysis. In *Inequality, Mobility and Segregation: Essays in Honor of Jacques Silber*. Bentley: Emerald Group Publishing Limited, pp. 135–59.
Esteban, Joan, Carlos Gradín, and Debraj Ray. 2007. An extension of a measure of polarization, with an application to the income distribution of five OECD countries. *The Journal of Economic Inequality* 5: 1–19. [CrossRef]
Esteban, Joan-Maria, and Debraj Ray. 1994. On the Measurement of Polarization. *Econometrica: Journal of the Econometric Society* 62: 819–51. [CrossRef]
Esteban, Joan-Maria, and Debraj Ray. 1999. Conflict and distribution. *Journal of Economic Theory* 87: 379–415. [CrossRef]
Esteban, Joan-Maria, and Debraj Ray. 2008. Polarization, fractionalization and conflict. *Journal of Peace Research* 45: 163–82. [CrossRef]
Esteban, Joan-Maria, and Debraj Ray. 2011. Linking conflict to inequality and polarization. *American Economic Review* 101: 1345–74. [CrossRef]
Ezcurra, Roberto. 2009. Does Income Polarization Affect Economic Growth? the Case of the European Regions. *Regional Studies* 43: 267–85. [CrossRef]
Firpo, Sergio, Nicole M. Fortin, and Thomas Lemieux. 2009. Unconditional Quantile Regressions. *Econometrica* 77: 953–73.
Firpo, Sergio P., Nicole M. Fortin, and Thomas Lemieux. 2018. Decomposing Wage Distributions Using Recentered Inuence Function Regressions. *Econometrics* 6: 28. [CrossRef]
Foster, James E., and Michael C. Wolfson. 1992. *Polarization and the Decline of the Middle Class: Canada and the US*. Working Paper 31. Oxford: Oxford Poverty & Human Development Initiative. Available online: https://ophi.org.uk/working-paper-number-31/ (accessed on 10 March 2023).
Foster, James E., and Michael C. Wolfson. 2010. Polarization and the Decline of the Middle Class: Canada and the US. *The Journal of Economic Inequality* 8: 247–73. [CrossRef]
Gasparini, Leonardo, Matias Horenstein, Ezequiel Molina, and Sergio Olivieri. 2008. Income polarization in Latin America: Patterns and links with institutions and conflict. *Oxford Development Studies* 36: 461–84. [CrossRef]
Gigliarano, Chiara, and Karl Mosler. 2009. Constructing indices of multivariate polarization. *The Journal of Economic Inequality* 7: 435–60. [CrossRef]
Gradín, Carlos. 2000. Polarization by Sub-Populations in Spain, 1973–1991. *Review of Income and Wealth* 46: 457–74. [CrossRef]
Handcock, Mark S., and Martina Morris. 1998. Relative Distribution Methods. *Sociological Methodology* 28: 53–97. [CrossRef]
Handcock, Mark S., and Martina Morris. 1999. *Relative Distribution Methods in the Social Sciences*. New York: Springer.
Hussain, Mohammad Azhar. 2009. The sensitivity of income polarization: Time, length of accounting periods, equivalence scales, and income definitions. *The Journal of Economic Inequality* 7: 207–23. [CrossRef]
Jann, Ben. 2021. Relative distribution analysis in Stata. *The Stata Journal* 21: 885–951. [CrossRef]
Keefer, Philip, and Stephen Knack. 2002. Polarization, Politics and Property Rights: Links Between Inequality and Growth. *Public Choice* 111: 127–54. [CrossRef]
Massari, Riccardo, Maria Grazia Pittau, and Roberto Zelli. 2009a. A Dwindling Middle Class? Italian Evidence in the 2000s. *Journal of Economic Inequality* 7: 333–50. [CrossRef]
Massari, Riccardo, Maria Grazia Pittau, and Roberto Zelli. 2009b. Caos calmo: L'evoluzione dei redditi familiari in Italia. In *L'Italia delle disuguaglianze*. Edited by Lorenzo Cappellari, Paolo Naticchioni and Stefano Staffolani. Rome: Carocci editore, pp. 19–28.
Milanovic, Branko. 2000. The median-voter hypothesis, income inequality, and income redistribution: An empirical test with the required data. *European Journal of Political Economy* 16: 367–410. [CrossRef]
Molini, Vasco, and Pierella Paci. 2015. *Poverty Reduction in Ghana—Progress and Challenges*. Washington, DC: World Bank.

Montalvo, José, and Marta Reynal-Querol. 2002. *Why Ethnic Fractionalization? Polarization, Ethnic Conflict and Growth*. UPF Economics and Business Working Paper No. 660. Barcelona: UPF Economics and Business.
Motiram, Sripad, and Nayantara Sarma. 2014. Polarization, Inequality, and Growth: The Indian Experience. *Oxford Development Studies* 42: 297–318. [CrossRef]
Nissanov, Zoya, and Maria Grazia Pittau. 2016. Measuring changes in the Russian middle class between 1992 and 2008: A nonparametric distributional analysis. *Empirical Economics* 50: 503–30. [CrossRef]
Petrarca, Ilaria, and Roberto Ricciuti. 2016. Relative income distribution in six European countries. In *Inequality after the 20th Century: Papers from the Sixth ECINEQ Meeting*. Bingley: Emerald Group Publishing Limited, vol. 24, pp. 361–86.
Pérez, Cristina Blanco, and Xavier Ramos. 2010. Polarization and Health. *Review of Income and Wealth* 56: 171–85. [CrossRef]
Poggi, Ambra, and Jacques Silber. 2010. On polarization and mobility: A look at polarization in the wage–career profile in Italy. *Review of Income and Wealth* 56: 123–40. [CrossRef]
Pressman, Steven. 2007. The Decline of the Middle Class: An International Perspective. *Journal of Economic Issues* 41: 181–200. [CrossRef]
Ricci, Chiara Assunta, and Sergio Scicchitano. 2021. Decomposing changes in income polarization by population group: What happened during the crisis? *Economia Politica* 38: 235–59. [CrossRef]
Rios-Avila, Fernando. 2020. Recentered influence functions (RIFs) in Stata: RIF regression and RIF decomposition. *The Stata Journal* 20: 51–94. [CrossRef]
Wolfson, Michael C. 1994. When inequalities diverge. *The American Economic Review* 84: 353–58.
Wolfson, Michael C. 1997. Divergent inequalities: Theory and empirical results. *Review of Income and Wealth* 43: 401–21. [CrossRef]
Zhang, Xiaobo, and Ravi Kanbur. 2001. What Difference Do Polarisation Measures Make? An Application to China. *Journal of Development Studies* 37: 85–98. [CrossRef]

economies

MDPI

Communication

Wage Inequality's Decreasing Effect on Enterprise Operating Revenues

Jarle Aarstad * and Olav Andreas Kvitastein

The Mohn Centre for Innovation and Regional Development, Western Norway University of Applied Sciences, P.O. Box 7030, NO-5020 Bergen, Norway; olav.andreas.kvitastein@hvl.no
* Correspondence: jarle.aarstad@hvl.no

Abstract: This study assesses whether wage inequality affects enterprises' operating revenues and whether operating revenues reversely affect wage inequality. To study our research questions, we analyze panel data from Norway and find that wage inequality decreases operating revenues. I.e., increasing high earners' wages relative to those earning low ones—or decreasing low earners' wages relative to those earning high ones—decreases operating revenues. It implies that wage inequality is detrimental to enterprise performance. Reversely, decreasing operating revenues increases wage inequality. I.e., low earners' wages are reduced relatively more than those earning high ones when enterprise revenues decrease. Increasing operating revenues, on the other hand, does not decrease wage inequality.

Keywords: dynamic unconditional quasi-maximum likelihood panel regression; dynamic GMM panel regression; instrumental variables

Citation: Aarstad, Jarle, and Olav Andreas Kvitastein. 2023. Wage Inequality's Decreasing Effect on Enterprise Operating Revenues. *Economies* 11: 178. https://doi.org/10.3390/economies11070178

Academic Editor: Fabio Clementi

Received: 28 April 2023
Revised: 26 June 2023
Accepted: 28 June 2023
Published: 30 June 2023

1. Introduction

This study aims to assess whether wage inequality affects enterprises' operating revenues and whether operating revenues reversely affect wage inequality. To study our research questions, we analyze a panel of more than 5000 Norwegian enterprises between 2008 and 2014. Largely we will argue in the following paragraphs that wage inequality is likely to decrease operating revenues, and reversely, we will argue that operating revenues are likely to increase wage inequality.

Wage inequality implies that a few employees earn relatively high wages compared to many others earning relatively low ones. Aarstad and Kvitastein (2021a) have summarized extensive literature examining wage inequality and shown that it is more prevalent in large rather than small industries in the number of enterprises (Aarstad and Kvitastein 2021b). In line with this research, other studies have found that profitability in the national economy, internal labor markets, and international trade affect wage inequality (Elgin et al. 2020; Nogueira and Afonso 2019; Pedace 2010).

An argument for wage inequality at an enterprise level is that high-earners, e.g., managers, align their interests with the owners (Beatty and Zajac 1994). This alignment of interests, in turn, induces the managers to increase enterprise revenues. High wages may moreover attract competent managers and other employees in key positions. Research has nonetheless indicated weak or absent performance effects from offering high wages to a small group of employees (Jensen and Murphy 1990; Kerr and Bettis 1987). Moreover, national-level research has even shown negative effects of economic inequality on growth and development (Berg et al. 2018; Voitchovsky 2005).

In this study, we similarly argue that wage inequality may have a negative effect on value creation. Taking an enterprise-level of analysis, we particularly suggest that wage inequality will decrease operating revenues. The reason for our assumption is that wage inequality implies that employees in lower ranks earn less than otherwise. Consequently, wage inequality may induce a sense of unfairness among most employees earning relatively

low salaries and hamper motivation. Taken together, we argue that wage inequality likely induces a lower production of goods and services, or lower quality of goods and services produced, which hampers operating revenues.

A parallel way to craft the argument is to assume an enterprise's wage budget as a fixed unit. I.e., an enterprise can decide to distribute the wage budget relatively evenly among the employees, which induces low wage inequality, or unevenly, which induces high wage inequality. In the latter case, a large share of the employees will earn less than in the former, which may hamper motivation. Consistent with our argument above, this issue precludes value creation and depresses operating revenues.

Having argued that wage inequality will likely negatively affect operating revenues, we do not rule out an opposite positive outcome. An argument for this is that wage inequality reflects key personnel, e.g., the management group and other critical employees, being particularly stimulated to increase their achievements on behalf of the enterprise, which in turn will increase operating revenues. Voitchovsky (2005) found that economic inequality among high-earners increases national growth and development, giving weight to this reasoning. Shortly, we test the two opposing arguments empirically.

Reversely, we study if operating revenues affect wage inequality. Operating revenues likely affect wages, but it is unclear how they affect their distribution. Yet having said that, Elgin et al. (2020) found that national profitability increased wage inequality, and Aarstad and Kvitastein (2021a) showed that industry-level operating profits increased wage inequality. In line with these studies, we assume that operating revenues will not only increase average wages but also increase wage inequality. I.e., those earning most at the outset will receive the largest relative wage premium when enterprise monetary resources abound.

2. Methodology

As noted, we study a panel of Norwegian enterprises between 2008 and 2014, and it is modeled by merging person-level data with enterprise-level data. Employees at year t were identified as those working full-time in the same enterprise at year t and t_{-1}. We included enterprises with at least 20 employees in the first year registered in the data. In the following years, we included observations of those same enterprises if they had at least ten employees. Enterprises with operations at more than one plant were excluded to avoid noise in the data concerning mergers, acquisitions, and demergers.

Our variables of primary interest are operating revenues and wage inequality at an enterprise level. In addition, we include average wages and enterprise size in full-time employees as control variables. The continuous variables were log-transformed, and Table 1 reports how they were measured.

Table 1. Variables.

Variable	Description
Operating revenues	Measured in 2014 prices by using Statistics Norway's consumer price index inflator.
Wage inequality	Gini index of full-time employees' wages.
Average wages	Based on full-time employees and measured in 2014 prices using Statistics Norway's wage index inflator.
Full-time employees	Counted straightforwardly.

3. Results

Table 2 reports dynamic unconditional quasi-maximum likelihood fixed-effects panel regressions with robust standard errors (Kripfganz 2016) and models independent and control variables at t and t_{-1}. Also, it includes unreported year dummies as controls (which is also the case in later tables).

Table 2. Dynamic unconditional quasi-maximum likelihood fixed-effects panels with robust standard errors.

	Model 1	**Model 2**	**Model 3**	**Model 4**
Dependent variable at t	Operating revenues		Wage inequality	
Dependent variable at t_{-1}	0.392 ***	0.392 ***	0.437 ***	0.436 ***
	(0.073)	(0.073)	(0.021)	(0.021)
Wage inequality at t	−0.066 **	−0.063 **		
	(0.022)	(0.021)		
Wage inequality at t_{-1}	0.027			
	(0.021)			
Operating revenues at t			−0.024 **	−0.022 **
			(0.009)	(0.008)
Operating revenues at t_{-1}			0.015 †	
			(0.008)	
Average wages at t	0.659 ***	0.661 ***	0.257 ***	0.261 ***
	(0.070)	(0.070)	(0.050)	(0.050)
Average wages at t_{-1}	−0.175 *	−0.171 †	0.012	0.020
	(0.087)	(0.088)	(0.047)	(0.046)
Full−time employees at t	0.451 ***	0.450 ***	0.097 ***	0.101 ***
	(0.047)	(0.047)	(0.015)	(0.015)
Full−time employees at t_{-1}	−0.104 ***	−0.102 ***	−0.047 ***	−0.041 ***
	(0.029)	(0.028)	(0.011)	(0.011)
Year dummies included	Yes	Yes	Yes	Yes
N enterprise−year obs./enterprises	20,082/5149	20,082/5149	20,082/5149	20,082/5149
Min./avg./max. obs. per enterprise	2/3.90/5	2/3.90/5	2/3.90/5	2/3.90/5

† $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Two-tailed tests for regressors and robust standard errors in parentheses.

Model 1 shows that a one percent increase in wage inequality at t significantly decreases operating revenues by 0.066 percent, but the effect at t_{-1} is non-significant. Omitting wage inequality at t_{-1} in Model 2 does not alter any statistical conclusion. Models 3 and 4 show that operating revenues at t have a significant negative effect on wage inequality, but the effect is not as marked as in the two previous models. Also, the coefficient is borderline significant positive at t_{-1} (Model 3), which indicates a weak "bounce back" effect.

Concerning the control variables, average wages seem to increase operating revenues, albeit leveling off somewhat the following year (Models 1 and 2). The probable reason for the increase is that wages stimulate the production of goods and services, but wages can also be a proxy for highly motivated and productive employees (theoretically, we cannot rule out reverse causality, but Table 3, using instrumental variables, concludes similarly). Unsurprisingly, increasing employment increases operating revenues, but the effect abates somewhat the following year (Models 1 and 2). Increasing average wages increases wage inequality (Models 3 and 4), which implies that those earning most at the outset take out a premium when overall wages increase. Finally, increasing employment increases wage inequality, albeit leveling off the following year (Models 3 and 4). The increase is probably because many newly recruited employees have relatively low experience and hence earn relatively low wages.

Table 3. Dynamic two-step Arellano-Bover/Blundell-Bond GMM panels with instrumental variables and robust standard errors.

	Model 1	**Model 2**	**Model 3**	**Model 4**
Dependent variable at t	Operating revenues		Wage inequality	
Dependent variable at t_{-1}	0.195	0.185	0.506 ***	0.498 ***
	(0.184)	(0.171)	(0.133)	(0.130)
Dependent variable at t_{-2}	0.126 *	0.130 *	0.047	0.051
	(0.056)	(0.051)	(0.060)	(0.059)
Wage inequality at t	−0.099 **	−0.103 **		
	(0.037)	(0.034)		
Wage inequality at t_{-1}	0.005			
	(0.022)			
Operating revenues at t			−0.045 ***	−0.045 ***
			(0.011)	(0.011)
Operating revenues at t_{-1}			0.004	
			(0.012)	
Average wages at t	1.13 ***	1.13 ***	0.286 ***	0.288 ***
	(0.179)	(0.174)	(0.059)	(0.059)
Average wages at t_{-1}	−0.117	−0.115	0.044	0.046
	(0.117)	(0.115)	(0.053)	(0.050)
Full−time employees at t	0.750 ***	0.753 ***	0.108 ***	0.111 ***
	(0.088)	(0.083)	(0.017)	(0.015)
Full−time employees at t_{-1}	−0.063	−0.061	−0.042 *	−0.040 *
	(0.048)	(0.045)	(0.017)	(0.015)
Year dummies included	Yes	Yes	Yes	Yes
Wald χ^2	2072.7 ***	2.79×10^6 ***	395.8 ***	391.9 ***
Second order z−value [a]/p−value	−1.36/0.173	−1.51/0.131	−0.13/0.896	−0.207/0.845
Hansen J test of over−id./p−value	5.44/0.908	4.59/0.970	10.9/0.456	11.2/0.515
Diff−in−Hansen (exl. group)/p−value	4.45/0.955	3.14/0.925	9.03/0.251	9.58/0.296
Diff−in−Hansen (difference)/p−value	3.03/0.882	1.45/0.836	1.82/0.768	1.58/0.812
Number of instruments	27	27	27	27
N enterprise−year obs./enterprises	21,017/6018	21,017/6018	21,017/6018	21,017/6018
Min./avg./max. obs. per enterprise	1/3.49/5	1/3.49/5	1/3.49/5	1/3.49/5

† $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Two-tailed tests for regressors and robust standard errors in parentheses. [a] Arellano–Bond test for zero autocorrelation in first-differenced errors.

Table 3 replicates the previous analyses using the dynamic two-step Arellano-Bover/Blundell-Bond GMM panel regression with instrumental variables. Our motive for this estimation technique is that we cannot rule out independent variables as being strictly exogenous, i.e., we cannot rule out that "they are correlated with … possibly current realizations of the error [term]" (Roodman 2009, p. 86). In lay terms, it implies that we cannot rule out independent variables at year t being affected by the current values of the dependent variable (Li et al. 2021). Specifically concerning our study, we have argued that wage inequality affects operating revenues and that operating revenues reversely affect wage inequality, i.e., we have argued that the causality goes in both directions. We report heteroscedasticity bias-corrected (wc) robust standard errors (Arellano and Bover 1995; Blundell and Bond 1998; Windmeijer 2005). Also, we add the lagged dependent variable at t_{-2}. The statistical conclusions are largely unaltered, except that the independent variables at t show stronger effects and that the effect of operating revenues at t_{-1} on wage inequality is now non-significant.[1] The post-estimation autocorrelation tests and

the Hansen *J* overidentification tests are non-significant. Similarly, the post-estimation correlations between endogenous variables and unobserved fixed effects (two last tests) are non-significant, indicating valid instruments (for an explanation of these tests, please see, e.g., Li et al. 2021).

Model 1, Table 4, using fixed effects regressions with robust standard errors, replicates the second model in the two previous tables but omits the lagged dependent variable as it could otherwise have induced biased estimates (cf. Nickell 1981). Overall, the statistical conclusions are unaltered compared to the previous models. Model 2 (Model 3) only includes observations where wage inequality increases (decreases) from t to t_{-1}. Albeit Model 3 shows a borderline-significant effect, Models 2 and 3 indicate that increasing wage inequality decreases operating revenues while decreasing wage inequality increases them. Similar exercises in Models 4–6, switching the dependent and independent variable, show that increasing operating revenues do not decrease wage inequality (Model 5). Decreasing operating revenues, on the other hand, tend to increase it (Model 6).

Table 4. Dynamic Fixed-effects panels with robust standard errors.

	Model 1	**Model 2**	**Model 3**	**Model 4**	**Model 5**	**Model 6**
Dependent variable at t	Operating revenues			Wage inequality		
Wage inequality at t	−0.054 *	−0.075 *	−0.052 †			
	(0.021)	(0.035)	(0.027)			
Operating revenues at t				−0.019 *	−0.003	−0.023 †
				(0.007)	(0.012)	(0.014)
Average wages at t	0.740 ***	0.618 ***	0.974 ***	0.205 **	0.343 ***	0.060
	(0.090)	(0.166)	(0.092)	(0.071)	(0.055)	(0.126)
Average wages at t_{-1}	0.008	−0.061	−0.069	0.116 **	0.049	0.180 **
	(0.059)	(0.122)	(0.076)	(0.034)	(0.038)	(0.052)
Full−time employees at t	0.567 ***	0.609 ***	0.547 ***	0.089 ***	0.081 ***	0.095 ***
	(0.029)	(0.058)	(0.038)	(0.016)	(0.019)	(0.025)
Full−time employees at t_{-1}	0.059 **	0.009	0.084 **	−0.015	−0.030 *	0.001
	(0.021)	(0.053)	(0.031)	(0.010)	(0.013)	(0.016)
Year dummies included	Yes	Yes	Yes	Yes	Yes	Yes
N enterprise−year obs./enterprises	27,898/6751	14,047/6127	13,851/6075	27,898/6751	15,565/5996	12,333/5597
Min./avg./max. obs. per entreprise	1/4.1/6	1/2.3/6	1/2.3/6	1/4.1/6	1/2.6/6	1/2.2/6
F−value	151.6 ***	85.3 ***	69.2 ***	14.1 ***	8.22 ***	6.88 ***
R−sq. within/between	0.232/0.583	0.259/0.597	0.229/0.571	0.020/0.077	0.027/0.102	0.022/0.052
Wage inequality at $t > t_{-1}$		Yes				
Wage inequality at $t < t_{-1}$			Yes			
Operating revenues at $t > t_{-1}$					Yes	
Operating revenues at $t < t_{-1}$						Yes

† $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Two-tailed tests for regressors and robust standard errors in parentheses.

4. Discussion

This study shows that wage inequality decreases enterprises' operating revenues. I.e., increasing high earners' wages relative to those earning low ones—or decreasing low earners' wages relative to those earning high ones—decreases operating revenues. The finding implies that wage inequality is detrimental to enterprise performance, and an explanation may be that decreasing low earners' wages may induce a sense of unfairness, hampering motivation among the majority earning low salaries. In turn, these issues affect the production of goods or services or decrease the quality of goods and services produced, reducing operating revenues.

Our findings align with studies showing economic inequality's negative effects on growth and development (Berg et al. 2018; Voitchovsky 2005). Also, they align with other research which indicates weak or absent performance effects from top management compensation (Jensen and Murphy 1990; Kerr and Bettis 1987). It implies that motivating numerous employees in relatively low positions with relatively high wages is more important for generating operating revenues than offering high wages and compensations to a relatively small group of employees in key positions. In other words, our findings counter the argument that increasing high earners' wages increase performance by attracting highly competent key personnel and aligning the interests of those in key positions with the owners' interests (cf. Beatty and Zajac 1994).

Reversely, our study shows that decreasing operating revenues increase wage inequality, i.e., low earners' wages are reduced relatively more than those earning high ones when operating revenues decrease. Conversely, increasing operating revenues does not decrease wage inequality, i.e., increasing operating revenues does not increase low earners' wages relative to those earning high ones.

A limitation of the study is that it only investigates a single national context. The results may deviate elsewhere in different cultures, which future research should investigate. A further limitation is that the study did not investigate which factors may genuinely explain the associations between wage inequality and operating revenues that we discovered, and this is also another topic for future research to delve into. A final limitation is that the study only investigated the changes in operating revenues as a relatively crude performance measure, and we, therefore, encourage future research to consider other output indicators. Moreover, finding that decreasing operating revenues increase wage inequality counters our argument and previous research, i.e., while industry- and national-level research shows that wage inequality increases when monetary resources abound (Aarstad and Kvitastein 2021a; Elgin et al. 2020), our enterprise-level study shows that wage inequality increases when operating revenues decrease. Level issues, measurement issues, or the use of control variables may explain the discrepancy, which we encourage future research to investigate.

5. Conclusions and Policy Implications

This study assessed whether wage inequality affects enterprises' operating revenues and whether operating revenues reversely affect wage inequality. To study our research questions, we analyzed a panel of more than 5000 Norwegian enterprises between 2008 and 2014.

The data showed that wage inequality decreases enterprises' operating revenues. I.e., increasing high earners' wages relative to those earning low ones—or decreasing low earners' wages relative to those earning high ones—decreases operating revenues. Reversely, the data showed that decreasing operating revenues increases wage inequality. I.e., low earners' wages are reduced relatively more than those earning high ones when operating revenues decrease. Conversely, increasing operating revenues does not decrease wage inequality, i.e., increasing operating revenues does not increase low earners' wages relative to those earning high ones.

A policy implication concerning the decreasing effect of wage inequality on operating revenues is that stockholders should reconsider managers' and other high-earning employees' compensation compared to those earning less. Our finding indicates that a

relatively equal wage distribution benefits enterprises' operating revenues and hence aligns with the stockholders' interests. A policy implication concerning the increasing effect of decreasing operating revenues on wage inequality is that those earning low wages at the outset should be aware of the issue, paying close attention to how they are compensated in the unfortunate event of an enterprise downscaling its operations.

Author Contributions: Conceptualization, J.A.; methodology, J.A. and O.A.K.; software, J.A.; validation, J.A. and O.A.K.; formal analysis, J.A.; writing—original draft preparation, J.A. and O.A.K.; writing—review and editing, J.A. All authors have read and agreed to the published version of the manuscript.

Funding: This research has no external funding.

Informed Consent Statement: Not applicable.

Data Availability Statement: For requests concerning the data, please contact the corresponding author.

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

Note

[1] The Model 1 (Table 3) Stata code is xtabond2 l(0/2)y l(0/1)(x1 x2 x3) i.year, gmm(l2.y x1 x2 x3, lag(1 .) collapse) two robust where y is the dependent variable, x1 wage inequality, x2 average wages, x3 full time employees, and i.year year dummies (see Roodman 2009). Thus, independent variables are treated as endogenous at t and predetermined at t_{-1}. The Model 2 code is xtabond 2 l(0/2)y x1 l(0/1)(x2 x3) i.year, gmm(l2.y x1 x2 x3, lag(1 .) collapse) two robust. Models 3 and 4 use similar codes.

References

Aarstad, Jarle, and Olav A. Kvitastein. 2021a. Do Operating Profits Induce a Wage Premium Equally Shared among Employees Earning High or Low Incomes? *Economies* 9: 81. [CrossRef]

Aarstad, Jarle, and Olav A. Kvitastein. 2021b. Is Industry Size a Carrier for Wage Inequality? A Panel Study Addressing Independent Variables of Inherently Different Sizes across Units. *Journal of Risk and Financial Management* 14: 436. [CrossRef]

Arellano, Manuel, and Olympia Bover. 1995. Another look at the instrumental variable estimation of error-components models. *Journal of Econometrics* 68: 29–51. [CrossRef]

Beatty, Randolph P., and Edward J. Zajac. 1994. Managerial incentives, monitoring, and risk bearing: A study of executive compensation, ownership, and board structure in initial public offerings. *Administrative Science Quarterly* 39: 313–35. [CrossRef]

Berg, Andrew, Jonathan D. Ostry, Charalambos G. Tsangarides, and Yorbol Yakhshilikov. 2018. Redistribution, inequality, and growth: New evidence. *Journal of Economic Growth* 23: 259–305. [CrossRef]

Blundell, Richard, and Stephen Bond. 1998. Initial conditions and moment restrictions in dynamic panel data models. *Journal of Econometrics* 87: 115–43. [CrossRef]

Elgin, Ceyhun, Adem Yavuz Elveren, and Joseph Bourgeois. 2020. Informality, Inequality and Profit Rate. *Applied Economics Letters* 28: 1017–20. [CrossRef]

Jensen, Michael C., and Kevin J. Murphy. 1990. Performance pay and top management incentives. *Journal of Political Economy* 98: 225–64. [CrossRef]

Kerr, Jeffrey, and Richard A. Bettis. 1987. Boards of directors, top management compensation, and shareholder returns. *Academy of Management Journal* 30: 645–64. [CrossRef]

Kripfganz, Sebastian. 2016. Quasi–maximum likelihood estimation of linear dynamic short-T panel-data models. *The Stata Journal* 16: 1013–38. [CrossRef]

Li, Jiatao, Haoyuan Ding, Yichuan Hu, and Guoguang Wan. 2021. Dealing with dynamic endogeneity in international business research. *Journal of International Business Studies* 52: 339–62. [CrossRef]

Nickell, Stephen. 1981. Biases in Dynamic Models with Fixed Effects. *Econometrica* 49: 1417–26. [CrossRef]

Nogueira, Manuel Carlos, and Óscar Afonso. 2019. Engines of the Skill Premium in the Portuguese Economy. *CESifo Economic Studies* 65: 318–41. [CrossRef]

Pedace, Roberto. 2010. Firm Size-Wage Premiums: Using Employer Data to Unravel the Mystery. *Journal of Economic Issues* 44: 163–82. [CrossRef]

Roodman, David. 2009. How to do Xtabond2: An Introduction to Difference and System GMM in Stata. *The Stata Journal* 9: 86–136. [CrossRef]

Voitchovsky, Sarah. 2005. Does the profile of income inequality matter for economic growth? *Journal of Economic Growth* 10: 273–96. [CrossRef]
Windmeijer, Frank. 2005. A finite sample correction for the variance of linear efficient two-step GMM estimators. *Journal of Econometrics* 126: 25–51. [CrossRef]

economies

Article

The Impact of Recent Economic Crises on Income Inequality and the Risk of Poverty in Greece

George Petrakos [1,2], Konstantinos Rontos [3], Chara Vavoura [4,*] and Ioannis Vavouras [1]

[1] Department of Public Administration, Panteion University of Social and Political Sciences, 17671 Athens, Greece; petrakos@panteion.gr (G.P.); vavouras@panteion.gr (I.V.)
[2] Research Institute for Tourism, 10564 Athens, Greece
[3] Department of Sociology, University of the Aegean, 81100 Mytilene, Greece; k.rontos@soc.aegean.gr
[4] Department of Economics, University of Athens, 10559 Athens, Greece
* Correspondence: cvavoura@econ.uoa.gr

Abstract: We consider the impact of the two recent economic crises, one that resulted from the great recession of 2007–2009 and one following the COVID-19 pandemic, on income inequality and the risk of poverty in Greece. To this end, we also investigate the key macroeconomic variables affecting the Greek income distribution. We find that alternative measures of inequality and relative poverty have deteriorated during the years of crises, implying that the country's response to economic shocks has been particularly harmful for its more disadvantaged citizens. Regarding the variables affecting the income distribution, we show that income inequality and the risk of poverty increase with growth, implying that the benefits of growth, and burdens of recession, are distributed unequally among members of the Greek society. Moreover, inequality appears to increase with unemployment and decrease with the share of wages and salaries in total income. This finding highlights the importance of labour market regulations and workers' welfare for attaining equality. Finally, we provide evidence that, during pre-election periods, income inequality is reduced, meaning that the political budget cycles, which are characteristic of the Greek economy, generate gains in terms of equality, but these gains are short-lived and mainly benefit the middle bracket of the income distribution.

Keywords: inequality; Gini coefficient; income quintile share ratio; risk of poverty rate; income inequality; economic crises

JEL Classification: D31; D63; E02

Citation: Petrakos, George, Konstantinos Rontos, Chara Vavoura, and Ioannis Vavouras. 2023. The Impact of Recent Economic Crises on Income Inequality and the Risk of Poverty in Greece. *Economies* 11: 166. https://doi.org/10.3390/economies11060166

Academic Editor: Fabio Clementi

Received: 25 April 2023
Revised: 26 May 2023
Accepted: 5 June 2023
Published: 13 June 2023

1. Introduction

Economic crises do not affect all members of a given society equally but embody redistributive mechanisms: those that have a limited capacity to respond are more vulnerable and, therefore, suffer greater losses. From this idea has sprung a long line of empirical works investigating the effects of economic crises on income distribution and the risk of poverty. Their conclusion, albeit not unanimous, is that income inequality, poverty, and the risk of social exclusion tend to increase during periods of crisis (Melidis and Tzagkarakis 2022; Bodea et al. 2021). This result mainly materializes through the negative impact on income distribution of the austerity measures undertaken to confront economic slowdowns (De Beer 2012).

Since the great recession fueled by the financial crisis of 2007–2009, Greece has become a synonym for economic crisis and the Greek economy has been in a recession for over half of the last fifteen years. This less-than-ideal situation has been the result of two major crises that took place in the country. The first one followed the great recession and took place between 2011 and 2016. It is known in the literature as the great economic slump. During those years, Greece faced the devastating consequences of its macroeconomic and structural imbalances: unprecedentedly high unemployment rates (in 2013 the unemployment rate

amounted to 27.5% of the labour force) combined with generally negative economic growth rates (in 2011 the growth rate of the real Gross Domestic Product (GDP) was −10.1%). Throughout the entirety of the great economic slump, it was only in 2014 that the real GDP growth rate turned marginally positive (0.5%)[1].

This crisis was followed by a three-year period (2017–2019) of significant de-escalation of unemployment (from 21.5% to 17.3%) and, at the same time, the GDP growth rate started to increase from an annual rate of 1.1% to 1.8%. Then, in 2020 and 2021, the Greek economy was hit by its second recent crisis, this time due to the COVID-19 pandemic. In 2020, economic growth was down by an astonishing 9.0%, while in 2021 the GDP growth rate was up to 8.3%, signaling a significant recovery of the economy boosted largely by the support measures undertaken both at a national and at a European level.

The main objective of this paper is to investigate the redistributive effects of these two almost consecutive crises of 2011–2016 and 2020–2021 and to measure their impact on the income distribution and the risk of poverty. To this end, we also investigate the main macroeconomic variables that affect the distribution of personal income and the risk of poverty in Greece.

When considering the issue of income distribution, what matters is personal income distribution, which refers to the way income is distributed among members of society regardless of its origin. The most widely used index is the "Gini coefficient of equivalized disposable income". Its value ranges from 0% (absolute equality) to 100% (absolute inequality). Another common criterion is the "income quintile share ratio S80/S20", which measures the ratio of the disposable income of the richest 20% of a given population to that of the poorest 20%. Notice that the S80/S20 is affected by the extreme values of the income distribution, whereas the Gini coefficient is not affected by extreme values and, thus, arguably reflects inequalities in a more satisfactory manner.

When we consider the effect of crises on poverty, we investigate the effect of crises on relative poverty and, more specifically, on the "at-risk-of-poverty rate"[2]. This index, being defined as the share of people with an equivalized disposable income below the at-risk-of-poverty threshold, set at 60% of the national median equivalized disposable income after social transfers, is a measure of low incomes in comparison to that of other residents in the country. As a result, this indicator operates as another measure of income inequality.

The phenomena of income inequality and poverty in Greece have been the subject of extensive empirical analysis (See mainly Melidis and Tzagkarakis 2022; Kaplanoglou 2022; Kotsios 2022; Andriopoulou et al. 2018, 2019; Giannitsis and Zografakis 2018; Kaplanoglou and Rapanos 2018; Matsaganis and Leventi 2014; Mitrakos 2014; Koutsampelas and Tsakloglou 2013; Mitrakos and Tsakloglou 2012; Tsakloglou and Mitrakos 2006; Tsakloglou and Panopoulou 1998; Tsakloglou 1993, 1990). This paper is, to our knowledge, the first attempt to examine how the aforementioned issues have been affected by the recent economic crises. As such, it also adds to the extensive literature on crises and inequality, surveyed in Dabla-Norris et al. (2015), in a substantial and policy-relevant way and, we believe, our results could serve as guidelines for a more socially desirable response to future economic setbacks.

Our paper also contributes to the ongoing discussion around the key macroeconomic factors that could potentially influence income distribution and the risk of poverty[3]. More specifically, we investigate the effect of the unemployment rate, the real GDP growth rate, and the share of labour income (in its narrow sense, that is, wages and salaries) in the GDP. We go on to summarize the main findings on the relationship of these three variables with income inequality, as reported in a number of relevant empirical articles.

Firstly, the extent of unemployment and income inequality have been found to be positively correlated, meaning that higher unemployment rates tend to be associated with greater inequality (e.g., Petrakos et al. 2022). Secondly, there is an opaque link between income distribution and economic growth dependent on a number of factors, such as market conditions and the extent of the state's redistributive role. The critical factor is the effectiveness of redistributive measures, mainly social transfers, used by the state (Vavoura

and Vavouras 2022). In the case of the EU member states, income inequality has been found to be positively correlated with economic growth in the cases of the economically developed member states and negatively correlated in the cases of the developing ones (Jianu et al. 2021).

Finally, a third factor considered is the share of wages and salaries in total income. Intuition suggests that as the share of wages and salaries in total income increases, income distribution should improve. However, this relationship also depends on the form and the progressiveness of the tax system in place, as well as on the effectiveness of the social welfare system. Hence, the correlation between labour income shares, imperfectly measured by the share of wages and salaries (Krueger 1999) and income inequality, is not clear (Erauskin 2020). Greece is particularly interesting in terms of its labour income because it presents the lowest share of wages and salaries in the total GDP compared to all EU member states during the period 1995–2021, for which data are available.

The remainder of the paper is structured as follows: Section 2 introduces the methodology, Section 3 presents the results, which are discussed in Section 4, and Section 5 concludes.

2. Methodology

In our analysis, we study the longitudinal evolution of personal income distribution and risk of poverty in the case of Greece during the period 1995–2021. We focus on 1995–2021 because it is for those years only that reliable data on income distribution and poverty exist. The data are derived from the Statistics on Income and Living Conditions of Households (SILC), which are based on the Eurostat methodology (EU-SILC). Before 1995 there are no official data available regarding the distribution of income in Greece, neither from the Hellenic Statistical Authority (ELSTAT), nor from any other public body, with the exception of individual estimates of researchers on the subject.

A simple overview of the evolution of the Gini coefficient in Greece and the European Union (EU) during the years 1995–2021 is presented in Figure 1. We observe, first, that income inequality in Greece is consistently above the EU average throughout the 1995–2021 period and, second, that during the years of the two crises (2011–2016 and 2020–2021) the Greek income distribution shows signs of deterioration. The current paper aims to establish that the second observation holds: that crises have indeed increased income inequality and the risk of poverty in the case of Greece.

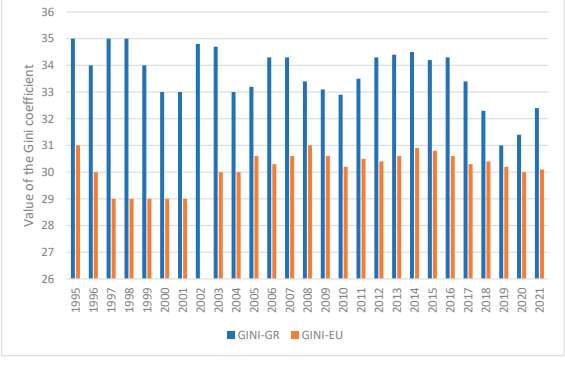


Figure 1. The evolution of the value of the Gini coefficient in Greece and the European Union during the period 1995–2021. Data sources: (1) Eurostat, EU Statistics on Income and Living Conditions (SILC), Gini coefficient of equivalized disposable income-EU-SILC survey. (2) ELSTAT, Statistics on Income and Living Conditions (SILC). Note: GINI-GR = the value of the Gini coefficient in Greece. GINI-EU = the average value of the Gini coefficient in the European Union.

It is important to mention that the pattern presented in Figure 1 is anything but universal. Focusing only on the EU, in the aftermath of the global crisis of 2007–2009, we

observe that between 2006 and 2009 the Gini coefficient of 12 member states increased, in 12 it decreased, and in 1 it remained unchanged, while for 2 there are no data available. Regarding the 2020–2021 COVID-19 pandemic, we observe that between 2019 and 2021 the income distribution, as measured by the value of the Gini coefficient, worsened in 9 EU member states, improved in 14, and in 3 it remained unchanged, while for 1 there are no data available. We argue that the effects of crises on income inequality vary with the policy mix that governments use to deal with them. In the case of the EU, the majority of the member states witnessed a betterment of their income distribution as a result of the two major recent crises[4]. To this rule, Greece, at first sight, appears to be an exception.

At this point, it is useful to explain why we have chosen to measure inequality via the Gini coefficient of disposable income, as opposed to the Gini coefficient of market income. In addition to data availability, since data for the Gini coefficient of market income is only available for the years 2006–2019[5], there is an additional methodological issue. By definition, disposable income is market income, that is, income from work and capital, after the deduction of direct taxes and payment of social security contributions. A comparison of market income and disposable income could be useful to evaluate the distributive role of the state. However, we argue that, when referring to income distribution and risk of poverty, what matters is the actual distribution of disposable income and not the income before the exercise of any redistributive role by the state. For this reason, the great bulk of the relevant empirical work relies on disposable income. Of course, the Gini coefficient of market income is always higher (depicting higher inequality) than the corresponding coefficient of disposable income, as the available data reveal. In the case of Greece for the period 2006–2019, for which data are available for market income, we can see this in Figure A4 of Appendix C.

To assess the robustness of our observations from Figure 1, we check our second measure of income inequality, which is the income quintile share ratio index S80/S20, and our measure of relative poverty, that is, the at-risk-of-poverty rate. A pattern similar to the pattern in Figure 1 is revealed: throughout the period under consideration, the values of the income quintile share ratio index S80/S20, presented in Figure 2, and the at-risk-of-poverty rate, shown in Figure 3, are systematically higher in Greece compared to the EU average. This implies that disparities between poor and rich incomes, as well as the percentages of the at-risk-of-poverty and social exclusion categories, have been consistently more severe in the Greek economy. More importantly, in the crises years under consideration (2011–2016 and 2020–2021), both indicators appear to deteriorate (see Figures 2 and 3). We proceed to formally test whether this second pattern can be verified. This takes us to the formal statement of our research hypothesis:

Hypothesis 1. *During the years of the two recent crises, income inequality and poverty have increased in Greece.*

To test this, we also investigate the most important macroeconomic variables that affect the intertemporal distribution of income in the country.

Our approach is an empirical one and it is based on the formulation and estimation of linear econometric models related to its subjects, which combine quantitative and dichotomous or dummy variables as independent variables. We construct three empirical models, each with one of the following three dependent variables:

(1) The Gini coefficient of equivalized disposable income (GINI), as defined and measured by the Eurostat Statistics on Income and Living Conditions (SILC) and the Hellenic Statistical Authority (ELSTAT).

(2) The income quintile share ratio index S80/S20 (QSR), as defined and measured by the Eurostat Statistics on Income and Living Conditions (SILC) and the Hellenic Statistical Authority (ELSTAT).

(3) The at-risk-of-poverty rate (ROPR), as defined and measured by the Eurostat Statistics on Income and Living Condition (SILC) and the Hellenic Statistical Authority (ELSTAT).

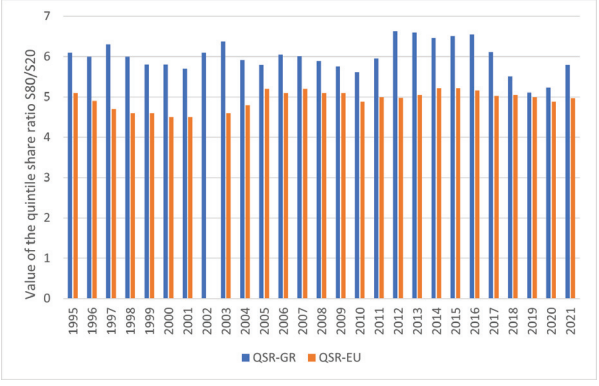


Figure 2. The evolution of the value of the income quintile share ratio S80/S20 index in Greece and the European Union during the period 1995–2021. Data sources: (1) Eurostat, EU Statistics on Income and Living Conditions (SILC), inequality of income distribution S80/S20 income quintile ratio, EU-SILC survey. (2) ELSTAT, Statistics on Income and Living Conditions (SILC). Note: QSR-GR = the value of the income quintile share ratio S80/S20 index in Greece. GINI-EU = the average value of the income quintile share ratio S80/S20 index in the European Union.

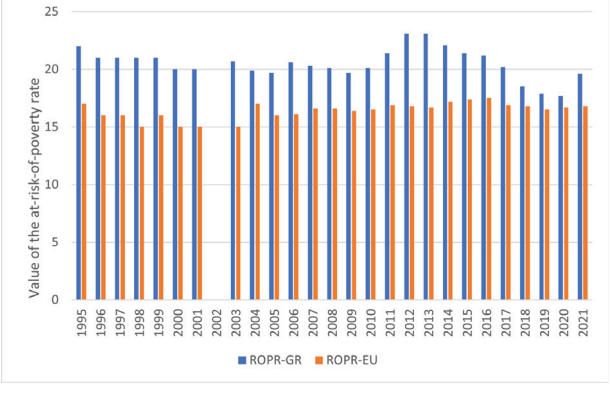


Figure 3. The evolution of the value of the at-risk-of-poverty rate in Greece and the European Union during the period 1995–2021. Data sources: (1) Eurostat, EU Statistics on Income and Living Conditions (SILC), at-risk-of-poverty rate by poverty thresholds, EU-SILC survey. (2) ELSTAT, Statistics on Income and Living Conditions (SILC). Note: ROPR-GR = the value of at-risk-of-poverty rate in Greece. ROPR-EU = the average value of at-risk-of- poverty rate in the European Union.

Based on the overview of the relevant theory and the purposes of our empirical analysis, we specify the following explanatory variables in the above three linear models. First, we use as explanatory variables the one-year lagged dependent variables to account for the fact that income inequality and poverty can have an autocorrelated AR(1) component and, hence, may be determined by past levels[6].

(4) The one-year lagged Gini coefficient (GINI-1) since the value of the current Gini coefficient can be influenced by its value in the previous year.

(5) The S80/S20 index with one year lag (QSR-1) since the value of the current S80/S20 coefficient may be affected by its value in the previous year.

(6) The at-risk-of-poverty rate with one year lag (ROPR-1) since the value of the current ROPR also has the potential to be shaped by its past value.

The second group of explanatory variables is related to the share of labour income, or rather dependent labour income, in total income. We include:

(7) The share of wages and salaries in GDP (WGDP), as measured by ELSTAT and the Eurostat methodology. This variable is particularly interesting in the case of Greece because the country is characterised by the lowest share of wages and salaries in the total GDP compared to all EU member states, during the period 1995–2021 for which data are available. Only three EU member states, and only for a short, specific time period, presented a lower share of wages and salaries in the GDP. Namely, Bulgaria during the period 1995–1997, Romania during the period 2009–2012, and Ireland during the period 2015–2021[7].

Our third group of explanatory variables is chosen so as to better capture the macroeconomic impact of an economic crisis. After the global financial crisis of 2007–2009, which developed into what is widely known as the great recession of 2007–2009, Greece entered a prolonged period of economic and fiscal crisis. Exit from this crisis became possible only in 2017. During that period, extensive literature (see, for example, Petrakos et al. 2022) has documented the following. Between 2008 and 2013 the unemployment rate jumped from 7.8% to an unprecedented 27.5% and, from there, it started to decrease gradually. Furthermore, during the years between 2008 and 2016, the country's total real GDP shrunk from €247.8 billion to €182 billion, a decrease corresponding to more than a quarter of its real value (26.6%). The most critical years of the 2008–2016 period were the years 2011–2016, which were characterised by unemployment rates higher than 20% and were, at the same time, associated with negative economic growth rates. Hence, we include in our models:

(8) The total real GDP in chainlinked volumes (2010) in millions of euros (RGDP) as measured by ELSTAT and the Eurostat methodology.

(9) The overall (total) annual unemployment rate (UNR) as measured by ELSTAT and the Eurostat methodology.

To address our research hypothesis on the impact of economic crises we are forced to include:

(10) A dummy variable for economic crises (GESCOV), which takes the value 1 during the years of crises in Greece (2011–2016 and 2020–2021) and 0 otherwise.

Finally, our last explanatory variable represents the impact of political factors. In particular, it captures the opportunistic behavior of governments during pre-election periods in order to win the upcoming general elections. This is the well-known phenomenon of "political electoral cycles" or "political budget cycles". The central hypothesis of this phenomenon is that incumbent politicians, faced with a "myopic" electorate with a decaying memory of past events (Bonfiglioli and Gancia 2013) or taking advantage of informational asymmetries that exist between them and rational constituents (Rogoff and Sibert 1988), choose to maximize their voting function instead of their social welfare function (Nordhaus 1975). Consequently, governments may pursue opportunistic expansionary fiscal policies before elections to appear competent and offer voters an illusion of economic prosperity (Petrakos et al. 2021b). This behavior may affect the distribution of income in the short run due to the economic benefits provided to the electorate to soften their impressions about the current governance of the country. The limited empirical research that exists on this issue concludes that income inequality actually improves during pre-election periods (Sever and Yucel 2021). To allow for the well-documented presence of political budget cycles (e.g., Petrakos et al. 2021b) in the Greek economy, we include:

(11) A dummy variable for elections (ELE), which takes the value 1 during the years of parliamentary elections in Greece during the period 1995–2021 and 0 otherwise.

Having specified our dependent and explanatory variables, we proceed to the specification of our empirical models. Let y_{mi} denote, for $m = 1, 2, 3$, respectively, the annual values of indices GINI, QSR, and ROPR, considered as dependent variables in the ith year, $i = 1, 2, \ldots, 27$. Let x_{ji}, $j = 1, 2, \ldots, 6$, denote the values of Xj corresponding to the

6 quantitative explanatory variables GINI-1, QSR-1, ROPR-1, WGDP, RGDP, and UNR, respectively. Let, in addition, z_{ki}, $k = 1, 2$ denote, respectively, the values of the two dummy variables GESCOV and ELE.

We assume that the mean value of the dependent variable of each sample can be expressed as a linear combination of the quantitative variables and the dummy variables in the following way:

$$E(y_{mi}) = \beta_0 + \sum_{j=1}^{6} \beta_j x_{ji} + \sum_{k=1}^{2} \gamma_k z_{ki}$$

We derive the following three models, which were tested against the usual assumptions of the OLS models:

- Model 1 ($m = 1$): dependent variable GINI.
- Model 2 ($m = 2$): dependent variable QSR.
- Model 3 ($m = 3$): dependent variable ROPR.

In order to access and compare the predictive accuracy of the three models, we use the Mean Absolute Percentage Error (*MAPE*) for $m = 1, 2, 3$:

$$MAPE(m) = \frac{1}{n} \sum_{i=1}^{n} \frac{|y_{mi} - \hat{y}_{mi}|}{|y_{mi}|} \times 100\%$$

3. Results

This section contains the results of our analysis. We present our results for Model 1, with GINI as the dependent variable, Model 2, with QSR as the dependent variable, and Model 3, with ROPR as the dependent variable.

3.1. Model 1 (m = 1): Dependent Variable GINI

In Model 1, the statistically significant independent variables are GINI-1, WGDP, RGDP, and the dummies GESCOV and ELE. The coefficient estimates and the significance tests are presented in Table 1 and the analysis of variance in Table A1 of Appendix A. Model 1 shows a coefficient of determination $R^2 = 71.12$ (R^2adj = 63.9). The Model 1 residuals are symmetrically distributed around zero with a standard deviation of 0.6281 and the largest one, corresponding to the eighth observation, is valued at 1.548. Nevertheless, there are not studentized residuals with Bonferroni $p < 0.05$ to be considered as outliers.

Table 1. Estimates of the coefficients of Model 1: dependent variable GINI.

Term	Coef.	SE Coef.	t-Value	*p*-Value	VIF
β_0	23.03	7.36	3.13	0.005	
GINI-1	0.481	0.165	2.92	0.008	1.90
WGDP	−0.322	0.108	−2.97	0.008	2.82
RGDP	0.000014	0.000006	2.18	0.041	1.59
GESCOV = 1	1.212	0.373	3.25	0.004	1.95
ELE = 1	−0.616	0.273	−2.26	0.035	1.04

From the Durbin–Watson test, the D-W Statistic is 2.27 with *p*-value at 0.966, larger than any significance level, concluding no autocorrelation at lag 1 in the model. Applying the studentized Breusch–Pagan test, we end up with BP statistics at 3.6664 on 5 df with *p*-value = 0.5984 and, hence, we cannot reject homoscedasticity of the model residuals. Finally, the Shapiro–Wilk test for normality with W statistic at 0.96019 and *p*-value = 0.3953 fails to reject the normality of the residuals. The above-mentioned test of the assumptions and relative conclusions is also illustrated in the diagnostics plots in Figure A1 of Appendix B. Searching for multicollinearity, the values of the VIF in Table 1, being less than three, indicate that there is no multicollinearity effect among the predictors.

In Model 1, the coefficient corresponding to WGDP has a negative sign, which indicates that increases in WGDP, i.e., increases in the share of wages and salaries in GDP, are

associated with decreases in the value of the GINI, implying a decrease in income inequality. In contrast, the GESCOV dummy variable comes in with a positive sign, which suggests that crises years are associated with increases in income inequality. At the same time, the dummy variable ELE has a negative coefficient, which suggests that during parliamentary election years, income inequality is reduced in the short term. This result is related to the phenomenon of political budget cycles. Furthermore, since RGDP enters with a positive sign, Model 1 confirms the hypothesis that income inequality is positively related to economic growth in the cases of developed countries. Finally, in the context of the present model, the unemployment rate (UNR) is not a statistically significant explanatory variable.

3.2. Model 2 (m = 2): Dependent Variable QSR

In Model 2, the statistically significant independent variables are QSR-1, WGDP, RGDP, and the dummy variable GESCOV. The coefficient estimates and significance tests of the model are presented in Table 2 and the analysis of variance in Table A2 of Appendix A. Model 2 shows a coefficient of determination R^2 = 75.79 (R^2adj = 71.17). The Model 2 residuals are positively skewed, with a standard deviation of 0.21 and the largest one, corresponding to the eighth observation, is valued at 2.17. Nevertheless, there are not studentized residuals with Bonferroni $p < 0.05$ to be considered as outliers.

Table 2. Estimates of the coefficients of Model 2: dependent variable QSR.

Term	Coef.	SE Coef.	t-Value	*p*-Value	VIF
β_0	4.46	1.16	3.86	0.001	
QSR-1	0.528	0.117	4.52	0.000	1.18
WGDP	−0.1057	0.0303	−3.48	0.002	1.95
RGDP	0.000005	0.000002	2.37	0.028	1.52
GESCOV = 1	0.634	0.127	4.98	0.000	2.00

From the Durbin–Watson test, the D-W Statistic is 2.11 with *p*-value at 0.572, larger than any significance level, concluding no autocorrelation at lag 1 in the model. Applying the studentized Breusch–Pagan test, we end up with BP statistics at 1.005 on 4 df with *p*-value = 0.909. Thus, we cannot reject the homoscedasticity of the model residuals. Finally, the Shapiro–Wilk test for normality, with W statistic at 0.8617 and *p*-value = 0.0024, rejects the normality of the residuals. The above-mentioned test of the assumptions and relative conclusions is also illustrated in the diagnostics plots in Figure A2 of Appendix B. Searching for multicollinearity, the values of the VIF in Table 2, being less than two, indicate that there is no multicollinearity effect among the predictors.

In Model 2, as in Model 1, the variable WGDP has a negative coefficient, which implies that increases in the share of wages and salaries in GDP are associated with decreases in income inequality, as measured by QSR. Furthermore, as in Model 1, the GESCOV dummy variable has a positive sign, which suggests that crises years are associated with increases in income inequality. As in Model 1, RGDP has a positive sign, meaning that income inequality is positively related to economic growth. In Model 2, the dummy variable ELE is not statistically significant, which suggests that the economic policy followed during the pre-election periods in the country, while limiting the overall inequality of income distribution as measured by the Gini coefficient, does not seem to affect the quintile share ratio index S80/S20 in a statistically significant way. That is, the ratio of the disposable income of the richest 20% of the population to that of the poorest 20% is not significantly affected. As in Model 1, the unemployment rate (UNR) is not a statistically significant explanatory variable.

3.3. Model 3 (m = 3): Dependent Variable ROPR

In Model 3, the statistically significant independent variables are WGDP, RGDP, UNR, and the dummy variable GESCOV. The coefficient estimates and significance tests of the model are presented in Table 3 and the analysis of variance in Table A3 of Appendix A.

Model 3 shows a coefficient of determination R^2 = 79.46 (R^2adj = 75.55). Model 3 residuals are symmetrically distributed around zero with a standard deviation of 0.65 and the largest one, corresponding to the eighth observation, is valued at 3.13. Nevertheless, there are not studentized residuals with Bonferroni $p < 0.05$ to be considered as outliers.

Table 3. Estimates of the coefficients of Model 3: dependent variable ROPR.

Term	Coef.	SE Coef.	t-Value	*p*-Value	VIF
β_0	32.15	1.90	16.93	0.000	
WGDP	−0.7339	0.0905	−8.11	0.000	2.04
RGDP	0.000027	0.000007	4.07	0.001	1.86
UNR	0.1268	0.0353	3.59	0.002	2.93
GESCOV = 1	1.769	0.456	3.88	0.001	2.71

From the Durbin–Watson test, the D-W Statistic is 1.44 with *p*-value at 0.02, concluding positive autocorrelation at lag 1 estimated at 0.15. Applying the studentized Breusch–Pagan test, we end up with BP statistics at 2.2322 on 4 df with *p*-value = 0.6931, meaning that we cannot reject the homoscedasticity of the model residuals. Finally, the Shapiro–Wilk test for normality, with W statistic at 0.949 and *p*-value = 0.22, fails to reject the normality of the residuals. The above-mentioned test of the assumptions and relative conclusions is also illustrated in the diagnostics plots in Figure A3 of Appendix B. Searching for multicollinearity, the values of the VIF in Table 3, being less than three, indicate that there is no multicollinearity effect among the predictors.

In order to cope with the autocorrelation detected with the D-W test, we introduce ROPR lag 1 (ROPR-1) and RORP lag 2 (ROPR-2) variables as predictors but get no significant improvement, since the lag variables are not significant, and autocorrelation remains at the same level.

In Model 3, as in Models 1 and 2, the variable WGDP has a negative coefficient, which implies that increases in the share of wages and salaries in GDP are associated with decreases in the at-risk-of-poverty rates. Crucially, as in Models 1 and 2, the GESCOV dummy variable has a positive coefficient, which suggests that crises years are associated with increases in the risk of poverty. As in Models 1 and 2, RGDP comes in with a positive sign, which confirms the hypothesis that the risk of poverty is positively related to economic growth in the cases of developed countries. Moreover, in Model 3 the unemployment rate (UNR) is a statistically significant explanatory variable, and its coefficient has a positive sign, which implies that increases in unemployment are associated with increases in the risk of poverty. Reducing unemployment can, therefore, be regarded as a significant means for reducing the at-risk-of-poverty rate. In Model 3, the dummy variable ELE is not statistically significant, which suggests that the economic policy followed during the pre-election periods in the country, while limiting the overall inequality of income distribution, does not seem to affect the risk of poverty in a statistically significant way.

3.4. Predictive Accuracy

In order to assess and compare the predictive accuracy of the three models, we use the Mean Absolute Percentage Error ($MAPE(m)$) with m = 1, 2, 3 for Models 1, 2, and 3, respectively. The results are: $MAPE(1) = 1.72$, $MAPE(2) = 2.86$, $MAPE(3) = 2.83$. Therefore, we conclude that Model 1 has better predictive accuracy when compared to the other two models.

3.5. Correlation between Inequality and Risk of Poverty

There is a significant positive correlation among all three dependent variables in the proposed regression models of the previous sections. The correlations were measured with the Pearson coefficient and are illustrated in Figure 4.

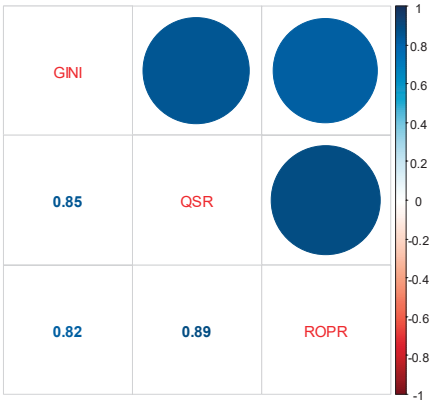


Figure 4. Correlations among dependent variables: Gini coefficient (GINI), quintile share ratio index S80/S20 (QSRR), and at-risk-of-poverty rate (ROPR).

By further investigating the correlation between income inequality as it is measured by the Gini coefficient and the risk of poverty in the case of Greece, we can conclude that:

1. Small values of the Gini coefficient GINI (low income inequality) are related to small values of the at-risk-of-poverty rate ROPR (low risk of poverty rate).
2. As illustrated in Figure 5, the lower values of income inequality (GINI) and risk of poverty (ROPR) are tied together in a close, linear relationship. However, in the higher value area for both variables, the relationship is somehow loose.

We can conclude that, in order to control the risk of poverty (ROPR), income inequality (GINI) has to be reduced.

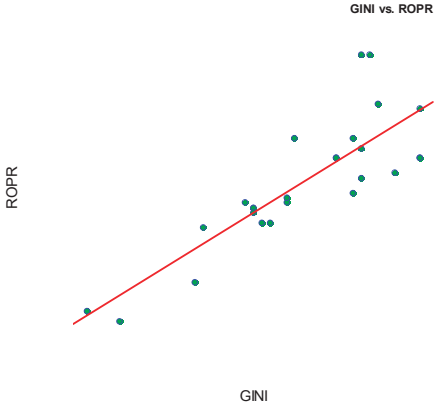


Figure 5. Gini coefficient (GINI) and at-risk-of-poverty rate (ROPR) xy plot.

4. Discussion

The aim of this paper is twofold. Firstly, we uncover the most important determinants of income inequality in Greece for the years 1995–2021. Secondly, we determine whether the economic crises that affected the Greek economy during 2011–2016, as a result of the global financial crisis, and during 2020–2021, as a result of the COVID-19 pandemic, improved or deteriorated income inequality and relative poverty in the country. Table 4 summarizes the results from our three empirical models presented in the previous section.

The first take-home point from our empirical analysis is that income inequality is autocorrelated. Regardless of the measure we use, income inequality depends heavily on past values like GINI-1 or QSR-1, and it appears that its adjustment process is slow: the one-year lagged inequality measure has a significant positive effect on the current measure,

meaning that a history of inequality creates further inequality. This result is supported by a number of different studies for different times and places. In fact, the persistence of high and, in many countries, rising income inequality over recent decades has become a growing concern at a global level (OECD 2008; Dabla-Norris et al. 2015).

Table 4. Summary of results.

Term	GINI	QSR	ROPR
GINI-1	0.481	-	-
QSR-1	-	0.528	-
WGDP	−0.322	−0.1057	−0.7339
RGDP	0.000014	0.000005	0.000027
UNR	-	-	0.1268
GESCOV = 1	1.212	0.634	1.769
ELE = 1	−0.616	-	-

The second factor of interest is the positive correlation between growth (RGDP) and inequality and poverty measures. This idea is not new in the literature, and it is actually common in the case of developed economies (Jianu et al. 2021). For Greece in particular, the relationship between income distribution and economic growth has been found to depend heavily on the effectiveness of the redistributive measures, mainly social transfers, used by the state (Vavoura and Vavouras 2022). We should not, however, dismiss the fact that the redistributive role of social transfers in Greece is limited to pensions, since all other forms of social transfers have been found to exert insignificant effects on income distribution (Papatheodorou and Petmesidou 2006). Consequently, the positive coefficient of the RGDP variable can be viewed as evidence of the limited redistributive role that Greek governments have played overtime, resulting in an unequal sharing of the benefits of economic growth (or an unequal sharing of the burdens of economic recession) among the members of Greek society.

This last point, regarding the role of the state, takes us to our next determinant of income inequality that appears to be instrumental in the case of Greece. We refer to the effect of the election (ELE) dummy, which is negatively correlated with inequality when we measure it via the Gini coefficient, but its effect becomes insignificant when we measure it via the quintile share ratio index S80/S20 or the at-risk-of-poverty rate. This result can be interpreted as follows. We know that the Greek economy is characterised by severe political budget cycles and that these cycles materialize mainly through an increase in government expenditure and, more specifically, via pre-election transfer payments (Petrakos et al. 2021a). These transfers decrease average income inequality, at least in the short term, but do not go as far as to affect extreme inequality (i.e., the share of the top and bottom 20%) and poverty. In other words, the pre-election transfers are mainly targeted towards the "average" voter in terms of personal income, possibly because it is precisely this type of citizen whose voting intentions are more volatile and, hence, more susceptible to opportunistic behavior on the side of the incumbent government. This idea, that governments distribute funding so as to influence the vote of swing voters, is, in fact, one that has been occasionally observed in European countries, for example, by Dahlberg and Johansson (2002) in Sweden and by Case (2001) in Albania.

Another important aspect of income inequality and poverty is their close ties to the labour market conditions. To begin with, a vast collection of literature documents the influence of unemployment on inequality and poverty (e.g., Mocan 1999; Cysne 2009; Zandi et al. 2022). In our study, unemployment has a significant positive correlation with the risk of poverty, although we do not find evidence that it significantly affects either of our other inequality measures. In any case, this result could help guide policy makers to increase their focus on unemployment when battling poverty and inequality. For Greece in particular, where the unemployment rate has not fallen below 15% during the period 2011–2021 according to Eurostat, we argue that any regulations aimed at the improvement

of income distribution are destined to fail unless they tackle the catastrophic economic and social consequences of prolonged unemployment.

However, the mechanism through which the labour market affects income inequality and poverty is not limited to unemployment. There is a second core factor: the labour income share. In our model, we follow the standard approach of measuring the labour income share via the share of wages and salaries in GDP (WGDP)[8]. This measure is known to have problems (Krueger 1999), mainly because it mismeasures the labour share of self-employment which, in Greece, is the highest among EU members[9]. Nonetheless, its strong negative correlation with all three measures of relative inequality and poverty (see Table 4), emphasizes the importance of workers' welfare and labour income protection as forces towards income equalization.

Related to the topic of income inequality, we should point out that there is a rapidly growing literature, following the works by Karabarbounis and Neiman (2014), Piketty and Zucman (2014), and Dao et al. (2017), on labour income share, documenting its worldwide decline. However, EU countries on aggregate have been found to follow no such trend. Cette et al. (2020) and Gutiérrez and Piton (2020) document the stability of the labour income share in Europe, with certain countries, Greece being one of them, witnessing an increase[10]. We argue that this trend has mitigated the effects of economic crises, which we are about to discuss, on Greek income distribution.

Finally, our most interesting result and our most notable contribution to the literature is summed up by the positive correlation between the crises (GESCOV) dummy and all three of our measures of inequality (GINI, QSR) and poverty (ROPR). This finding verifies our research hypothesis: the two recent economic crises were associated with a sharp increase in income inequality and relative poverty in Greece. The austerity measures adopted by EU institutions and the Greek government as a reaction to the country's debt crisis, fueled by the global financial crisis, as well as the policies that followed in response to the COVID-19 pandemic, operated as redistributive mechanisms, essentially "taking from the poor and giving to the rich". The redistributive role of the crises is stronger in the case of the lowest of incomes, given that the effect of the crises dummy is found to be higher in the at-risk-of-poverty rate. In other words, the crises in Greece have hit the economically weakest the hardest. Our results are in line with Nolan et al. (2019) who investigated the impact of the great recession of 2007–2009 and reported that Greece constitutes "the most dramatic case" of crisis-induced increased inequality.

To conclude, our study serves as a reminder that the central pilar of the Greek government's economic and social policy, especially during periods of crisis, should be to dampen their redistributive effects by increasing the role of the benefit and tax system in cushioning a declining aggregate income and in protecting households against income losses. However, to date, the policy mix adopted in order to deal with economic slowdowns does not appear to have contributed effectively to this direction.

Lastly, it should not go without mentioning that policy recommendations involving the Greek benefit and tax system can hardly be viewed as complete without mentioning the need to improve its effectiveness, to control hidden economy activity, and to restrict corruption (Manolas et al. 2013). Unless tax evasion, mainly among the incomes from entrepreneurial activities, is reduced, taxes and social security contributions will have but a weak impact on the reduction of inequality (Papatheodorou 1998).

5. Conclusions

The main objective of our analysis was to investigate whether the distribution of income and the risk of poverty in Greece in the post-1995 period were affected by the economic crises of the 2010s onward. We found that both income inequality and poverty worsened as a result of the crises and that this result is robust to different measures of inequality and poverty, namely the Gini coefficient, the quintile share ratio index S80/S20, and the at-risk-of poverty rate. All of these measures exhibit a significantly positive correlation with the economic crises' dichotomous variable, implying that the country's

response to economic shocks was particularly harmful for the economically disadvantaged members of the Greek society. We argue that this matter of extreme importance has received limited attention both from Greek economists and in public debate, and we hope that our study could help to shed light on the issue.

A second purpose of our work was to examine the most important macroeconomic variables affecting the distribution of income in Greece. Our findings can be summed up as follows. Firstly, increases in the share of wages and salaries in the GDP are associated with reductions in income inequality and the risk of poverty. Secondly, reductions of unemployment rates are associated with reductions of the risk of poverty. These two results suggest that labour market conditions are crucial for income inequality and that shielding workers' welfare is bound to increase equality, at least in the case of Greece. Thirdly, income inequality is positively related to economic growth. This result suggests that the benefits of growth and the burdens of recession are not distributed equally among the members of the Greek society. Finally, during pre-election periods, income inequality, as measured by the Gini coefficient, is reduced but extreme inequality, as measured by the quintile share ratio index S80/S20 and the at-risk-of poverty rate, is unaffected. This result implies that the political budget cycles that are characteristic of the Greek economy create gains in terms of equality, but these gains are short-lived and mainly benefit the middle bracket of the income distribution.

We conclude that the recent economic crises did not affect the Greek society in a balanced way, but rather had a strongly redistributive character. This does not, however, appear to have been the case for the majority of the rest of the EU member states. Consequently, we argue that there is a lesson to be learned for Greek policy makers who, in the future, should try to dampen these redistributive effects.

Future research, should such data become available, could consider how the effects of crises have varied for different regions of Greece and how they have depended on specific geographical and economic characteristics, possibly also taking into consideration tourism seasonality, as in Tsiotas et al. (2020), in order to propose useful regional policies.

Moreover, following works like Zeqiraj et al. (2022), Khan et al. (2014), and Kaasa (2003), principal component analysis could be applied to construct indicators of health and education, population and environmental conditions, sectoral growth and balance of payment factors, private investment, and public and social expenditures. Researching the link between such indicators and the income distribution could provide a spherical view of the evolution of income inequality in Greece and serve as the basis for the design of appropriate poverty combating policies.

Finally, from a cross-country study similar to Petrakos et al. (2023) of the four EU member states that were hit hardest by the global crisis of 2007–2009, namely Portugal, Italy, Greece, and Spain, it could be possible to draw interesting inferences regarding the impact of different austerity measures on the income distribution and to propose fewer equality-distorting alternatives for the future.

Author Contributions: Conceptualisation, I.V. and K.R.; methodology, G.P.; validation, G.P. and K.R.; formal analysis, G.P.; data curation, I.V.; writing—original draft preparation, C.V.; writing—review and editing, C.V. and I.V. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Publicly available datasets were analysed in this study. Data have been provided by Eurostat (ILC_DI12), EU Statistics on Income and Living Conditions (SILC) and the Hellenic Statistical Authority (ELSTAT).

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

Appendix A

This appendix contains the ANOVA tables for the regression models.

Table A1. Analysis of variance of Model 1: dependent variable GINI.

Source	DF	Adj. SS	Adj. MS	F-Value	*p*-Value
Regression	5	19.436	3.8871	9.85	0.000
GINI-1	1	3.366	3.3658	8.53	0.008
WGDP	1	3.489	3.4889	8.84	0.008
RGDP	1	1.881	1.8810	4.77	0.041
GESCOV = 1	1	4.174	4.1741	10.58	0.004
ELE = 1	1	2.014	2.0142	5.11	0.035
Error	20	7.891	0.3945		
Total	25	27.327			

Table A2. Analysis of variance of Model 2: dependent variable QSR.

Source	DF	Adj. SS	Adj. MS	F-Value	*p*-Value
Regression	4	2.9528	0.73820	16.43	0.000
QSR-1	1	0.9174	0.91743	20.42	0.000
WGDP	1	0.5454	0.54544	12.14	0.002
RGDP	1	0.2517	0.25174	5.60	0.028
GESCOV = 1	1	1.1144	1.11439	24.81	0.000
Error	21	0.9434	0.04492		
Total	25	3.8962			

Table A3. Analysis of variance of Model 3: dependent variable ROPR.

Source	DF	Adj. SS	Adj. MS	F-Value	*p*-Value
Regression	4	34.555	8.6387	20.31	0.000
WGDP	1	27.955	27.9547	65.73	0.000
RGDP	1	7.038	7.0379	16.55	0.001
UNR	1	5.491	5.4913	12.91	0.002
GESCOV = 1	1	6.400	6.3995	15.05	0.001
Error	21	8.932	0.4253		
Total	25	43.487			

Appendix B

This appendix contains the diagnostics plots for the regression models.

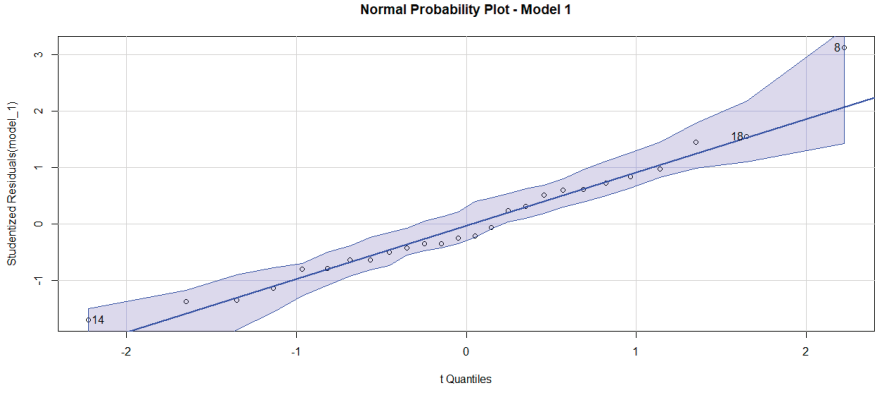


Figure A1. *Cont.*

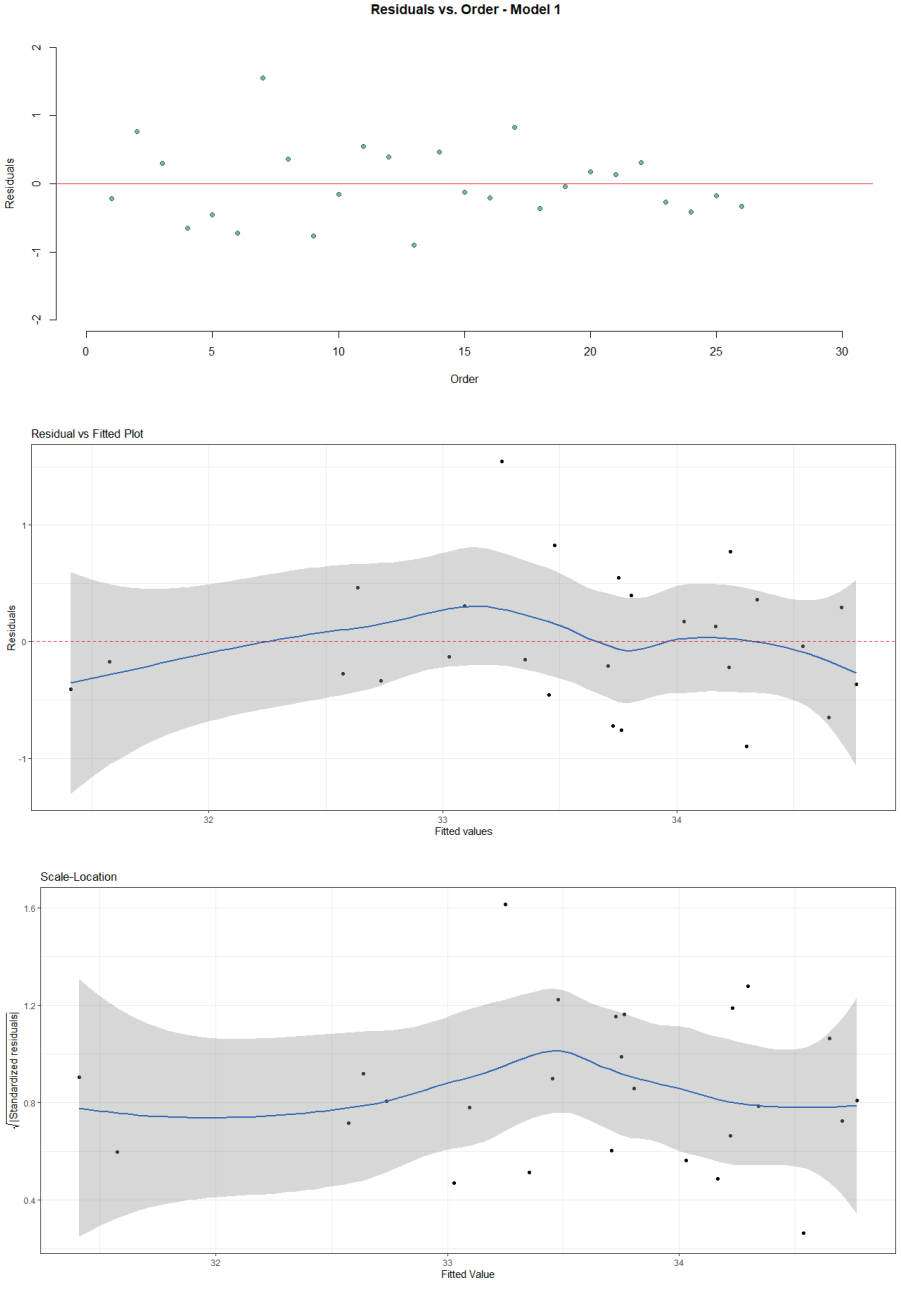


Figure A1. *Cont.*

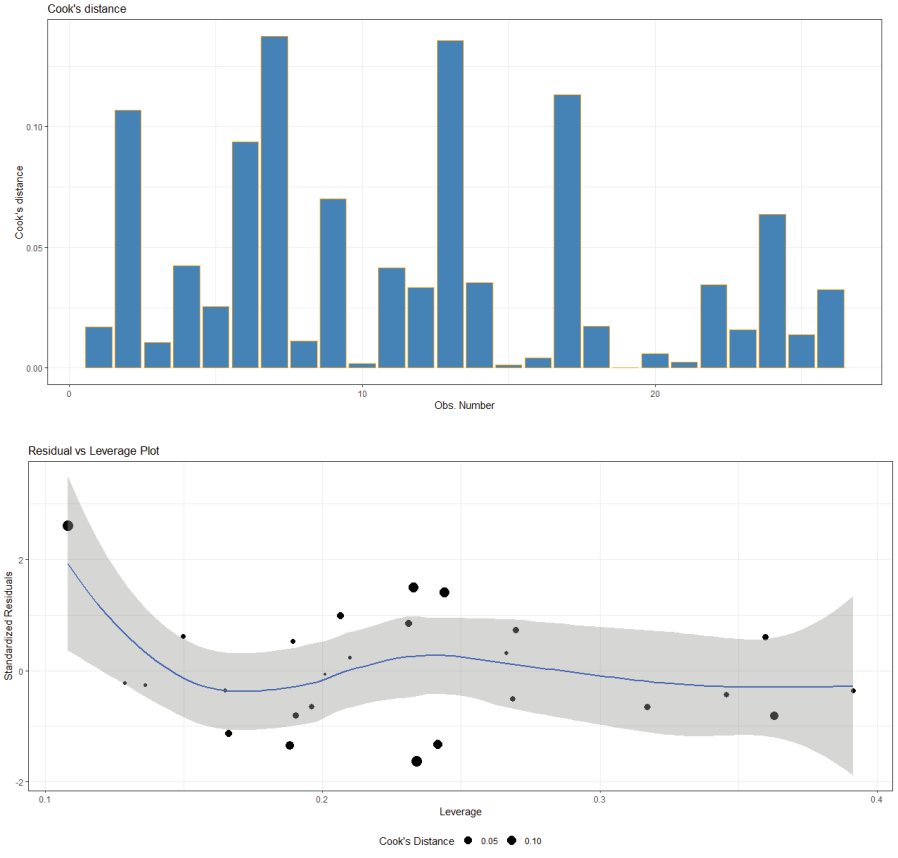


Figure A1. Diagnostic plots of Model 1: dependent variable GINI.

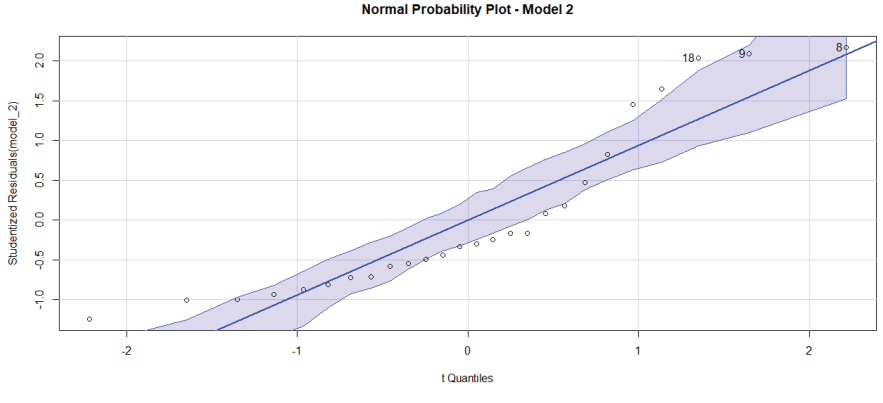


Figure A2. *Cont.*

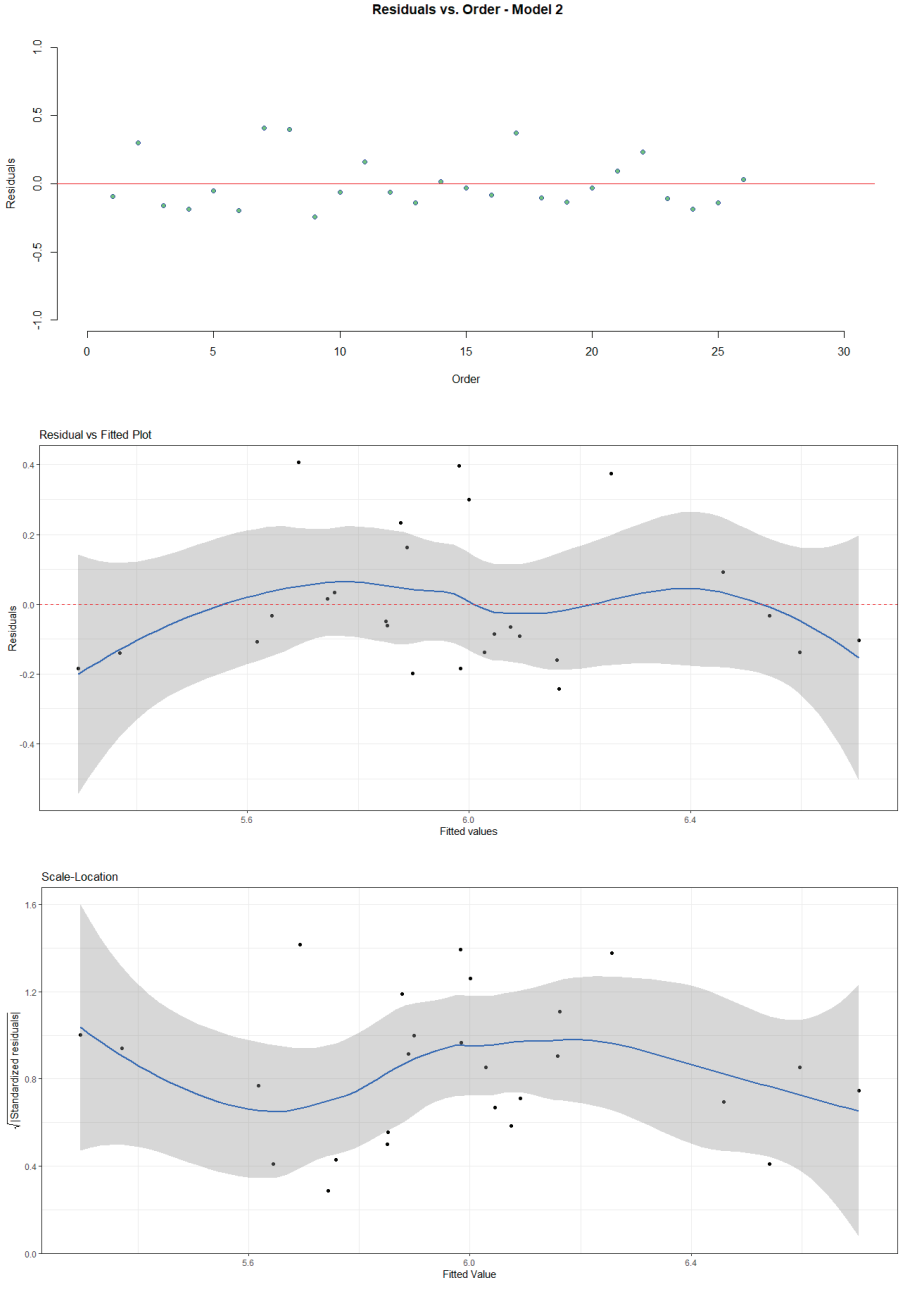


Figure A2. *Cont.*

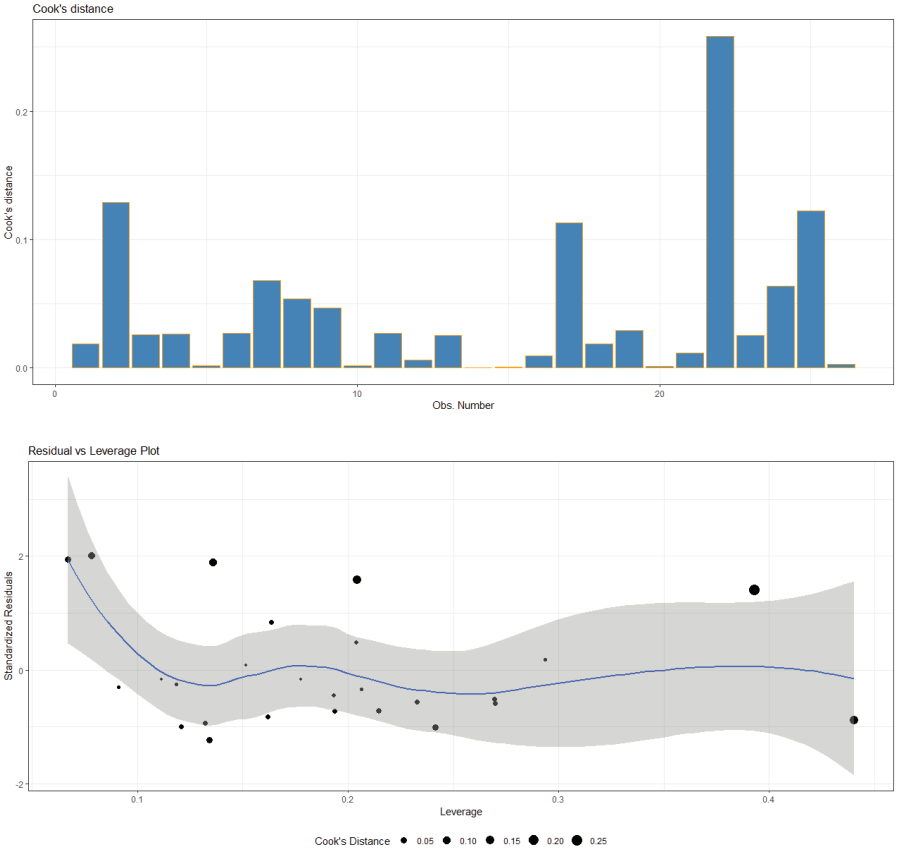


Figure A2. Diagnostic plots of Model 2: dependent variable QSR.

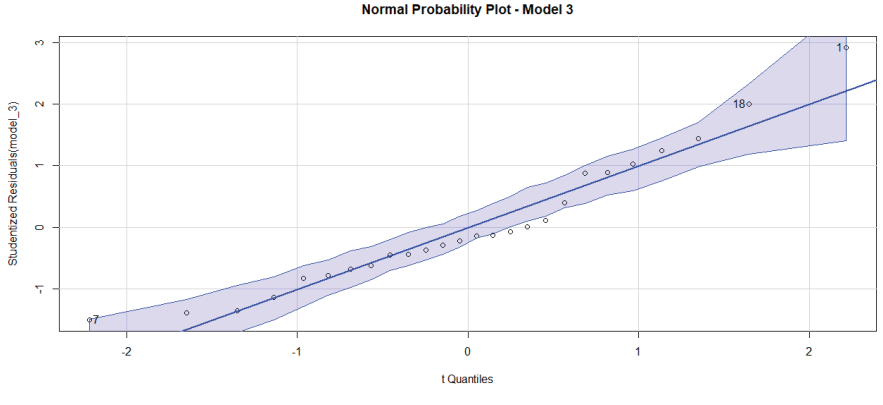


Figure A3. *Cont.*

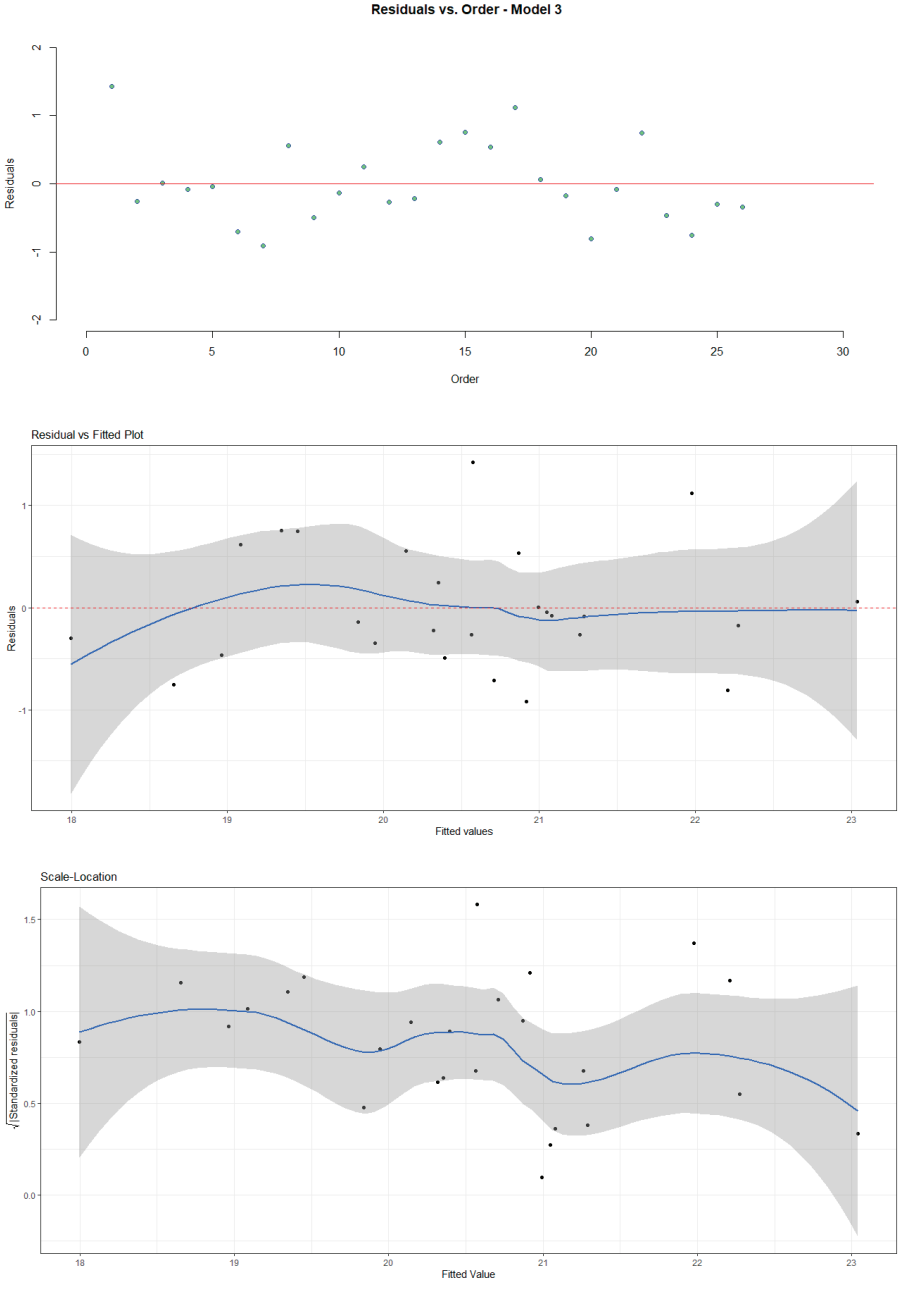


Figure A3. *Cont.*

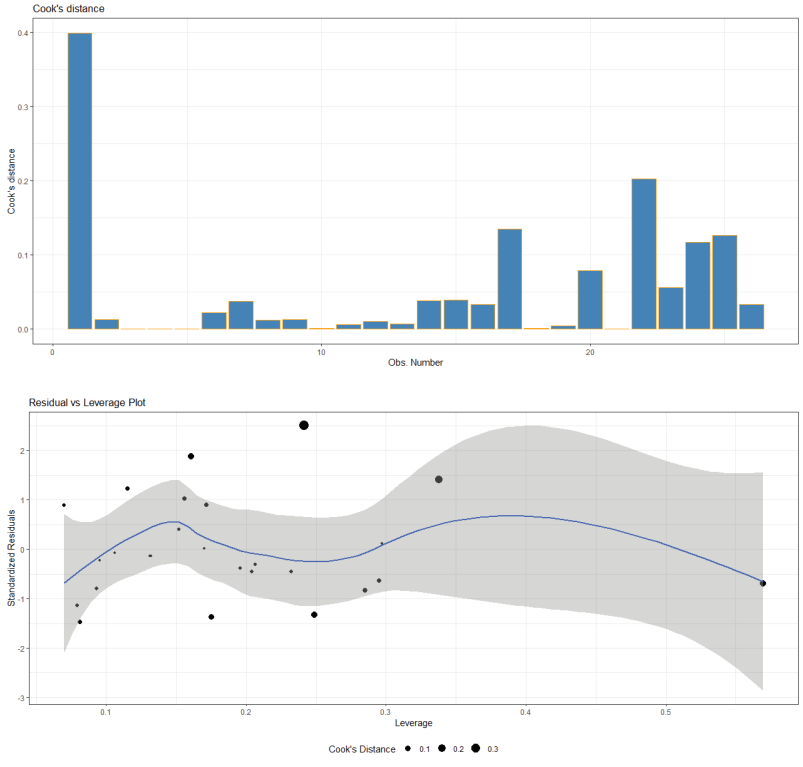


Figure A3. Diagnostic plots of Model 3: dependent variable ROPR.

Appendix C

This appendix contains the evolution of the Gini coefficient for market income and disposable income in Greece, for the years where data for both indices are available.

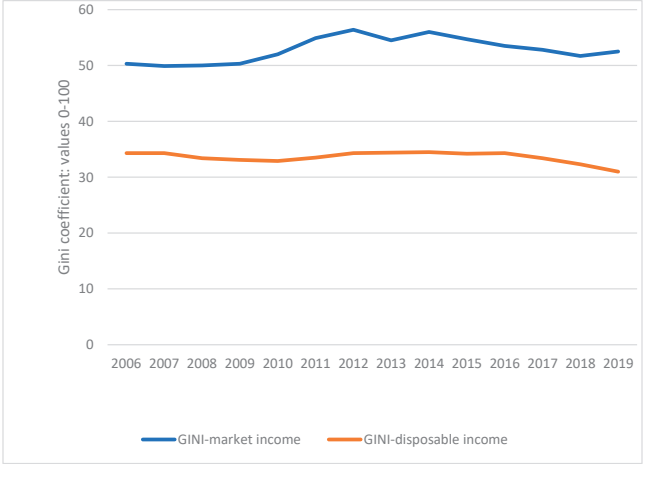


Figure A4. The evolution of the Gini coefficient for market income and disposable income in Greece during the period 2006–2019. Data sources: Gini coefficient for market income: OECD; Gini coefficient for disposable income: Eurostat; Gini coefficient of equivalized disposable income: EU-SILC survey (ilc-di12).

Notes

1 All data refer to Eurostat estimates.
2 See Eurostat (ILC_LI02), At-risk-of-poverty rate—EU-SILC and ECHP surveys. Online data code: TESPM010.
3 Microeconomic determinants of poverty and income inequality have also been found to be of major importance, particularly for the design of appropriate policies (see for example Naschold (2009)). However, in the case of Greece no such data are available for the years under consideration.
4 Eurostat (ILC_DI12), EU Statistics on Income and Living Conditions (SILC), Gini coefficient of equivalized disposable income—EU-SILC survey. Online data code: TESSI190.
5 Eurostat and ELSTAT provide estimates of the Gini coefficient of equivalized disposable income before social transfers (pensions included in social transfers) since 2003 (ilc-di12b) as well as estimates of the Gini coefficient of equivalized disposable income before social transfers (pensions excluded from social transfers) since 2003 (ilc-di12c). They do not provide estimates of the Gini coefficient based on market income. Only OECD, provides estimates of the Gini coefficient based on market income (before taxes and transfers) but only for the period 2006–2019.
6 We have also tested lagged variables of higher order and found no significant effect. Additionally, we run time series analysis for the y variables and also tried the Detrend model for which we have experienced violation of the basic assumptions. Given that our aim was to identify and quantify the relation of income inequality with certain predictors, we proceed and present the selected OLS models (with lag1 variable where appropriate).
7 Eurostat, GDP and main components (output, expenditure and income). Online data code: NAMA_10_GDP.
8 This is the narrow concept of labour income. The broad one is "compensation of employees" and it also includes employers' social contributions.
9 OECD (2023), Self-employment rate (indicator). doi: 10.1787/fb58715e-en.
10 Eurostat, GDP and main components (output, expenditure and income), nama_10_gdp.

References

Andriopoulou, Eirini, Alexandros Karakitsios, and Panos Tsakloglou. 2018. *Inequality and Poverty in Greece: Changes in Times of Crisis*. Cham: Springer International Publishing.
Andriopoulou, Eirini, Eleni Kanavitsa, and Panos Tsakloglou. 2019. Decomposing poverty in hard times: Greece 2007–2016. *Revue d' Économie du Développement* 27: 125–68. [CrossRef]
Bodea, Cristina, Christian Houle, and Hyunwoo Kim. 2021. Do financial crises increase income inequality? *World Development* 147: 105635. [CrossRef]
Bonfiglioli, Alessandra, and Gino Gancia. 2013. Uncertainty, electoral incentives and political myopia. *The Economic Journal* 123: 373–400. [CrossRef]
Case, Anne. 2001. Election goals and income redistribution: Recent evidence from Albania. *European Economic Review* 45: 405–23. [CrossRef]
Cette, Gilbert, Lorraine Koehl, and Thomas Philippon. 2020. Labor share. *Economics Letters* 188: 108979.
Cysne, Rubens Penha. 2009. On the positive correlation between income inequality and unemployment. *Review of Economics and Statistics* 91: 218–26. [CrossRef]
Dabla-Norris, Era, Kalpana Kochhar, Nujin Suphaphiphat, Frantisek Ricka, and Evridiki Tsounta. 2015. *Causes and Consequences of Income Inequality: A Global Perspective*. Bretton Woods: International Monetary Fund.
Dahlberg, Matz, and Eva Johansson. 2002. On the vote-purchasing behavior of incumbent governments. *American Political Science Review* 96: 27–40. [CrossRef]
Dao, Mai Chi, Mitali Das, Zsoka Koczan, and Weicheng Lian. 2017. *Why Is Labor Receiving a Smaller Share of Global Income? Theory and Empirical Evidence*. Bretton Woods: International Monetary Fund.
De Beer, Paul. 2012. Earnings and income inequality in the EU during the crisis. *International Labour Review* 151: 313–31. [CrossRef]
Erauskin, Iñaki. 2020. The labour share and income inequality: Some empirical evidence from the period 1990–2015. *Applied Economic Analysis* 28: 173–95. [CrossRef]
Giannitsis, Tassos, and Stavros Zografakis. 2018. *Crisis Management in Greece*. No. 50-2018. Bruxelles: IMK at the Hans-Böckler-Stiftung Foundation, Macroeconomic Policy Institute.
Gutiérrez, Germán, and Sophie Piton. 2020. Revisiting the global decline of the (non-housing) labor share. *American Economic Review: Insights* 2: 321–38. [CrossRef]
Jianu, Ionut, Marin Dinu, Dragos Huru, and Alexandru Bodislav. 2021. Examining the relationship between income inequality and growth from the perspective of EU member States' stage of development. *Sustainability* 13: 5204. [CrossRef]
Kaasa, Anneli. 2003. *Factors Influencing Income Inequality in Transition Economies*. University of Tartu Economics and Business Administration Working Paper Series, 18; Tartu: Faculty of Economics and Business Administration, University of Tartu.
Kaplanoglou, Georgia. 2022. Consumption inequality and poverty in Greece: Evidence and lessons from a decade-long crisis. *Economic Analysis and Policy* 75: 244–61. [CrossRef]

Kaplanoglou, Georgia, and Vassilis T. Rapanos. 2018. Evolutions in consumption inequality and poverty in Greece: The impact of the crisis and austerity policies. *Review of Income and Wealth* 64: 105–26. [CrossRef]
Karabarbounis, Loukas, and Brent Neiman. 2014. The global decline of the labor share. *The Quarterly Journal of Economics* 129: 61–103.
Khan, Muhammad Azhar, Muhammad Zahir Khan, Khalid Zaman, and Ayisha Khan. 2014. Poverty–growth–inequality triangle by principal component analysis: With an empirical illustration using Pakistan's data. *International Journal of Rural Management* 10: 69–86. [CrossRef]
Kotsios, Panagiotis. 2022. Income inequality measurement through tax data: The case of Greece. *Bulletin of Applied Economics* 9: 175–87. [CrossRef]
Koutsampelas, Christos, and Panos Tsakloglou. 2013. The distribution of full income in Greece. *International Journal of Social Economics* 40: 311–30. [CrossRef]
Krueger, Alan B. 1999. Measuring labour's share. *American Economic Review* 89: 45–51. [CrossRef]
Manolas, George, George Sfakianakis, and Ioannis Vavouras. 2013. Tackling the debt crisis in Greece: The role of the underground economy. *Social Cohesion and Development* 8: 25–35. [CrossRef]
Matsaganis, Manos, and Chrysa Leventi. 2014. Poverty and inequality during the Great Recession in Greece. *Political Studies Review* 12: 209–23. [CrossRef]
Melidis, Michail, and Stylianos I. Tzagkarakis. 2022. The evolution of social vulnerability in Greece during the economic crisis (2008–2017). *European Societies* 24: 229–50. [CrossRef]
Mitrakos, Theodore. 2014. *Inequality, Poverty and Social Welfare in Greece: Distributional Effects and Austerity*. Working Papers, No. 174. Athens: Bank of Greece.
Mitrakos, Theodore, and Panos Tsakloglou. 2012. Inequality and poverty in Greece: Myths, realities and crisis. In *Reforming Greece: Sisyphean Task or Herculean Challenge?* Edited by Othon Anastasakis and Dorian Singh. SEESOX Book Series on Current Affairs; Oxford: St Antony's College, pp. 90–99.
Mocan, H. Naci. 1999. Structural unemployment, cyclical unemployment, and income inequality. *Review of Economics and Statistics* 81: 122–34. [CrossRef]
Naschold, Felix. 2009. Microeconomic determinants of income inequality in rural Pakistan. *The Journal of Development Studies* 45: 746–68. [CrossRef]
Nolan, Brian, Matteo G. Richiardi, and Luis Valenzuela. 2019. The drivers of income inequality in rich countries. *Journal of Economic Surveys* 33: 1285–324. [CrossRef]
Nordhaus, William D. 1975. The political business cycle. *Review of Economic Studies* 42: 169–90. [CrossRef]
Organisation for Economic Co-operation and Development (OECD). 2008. *Growing Unequal? Income Distribution and Poverty in OECD Countries*. Paris: OECD Publishing.
Organisation for Economic Co-operation and Development (OECD). 2023. Self-Employment Rate (Indicator). Available online: https://www.oecd-ilibrary.org/employment/self-employment-rate/indicator/english_fb58715e-en (accessed on 6 June 2023).
Papatheodorou, Christos. 1998. *Inequality in Greece: An Analysis by Income Source*. LSE STICERD Research Paper 39. London: London School of Economics and Political Science.
Papatheodorou, Christos, and Maria Petmesidou. 2006. Poverty profiles and trends. How do southern European countries compare with each other? In *Poverty and Social Deprivation in the Mediterranean: Trends, Policies & Welfare Prospects in the New Millennium*. Edited by Maria Petmesidou and Christos Papatheodorou. London: Zed Books.
Petrakos, George, Konstantinos Rontos, Chara Vavoura, and Ioannis Vavouras. 2021a. The mechanism of political budget cycles in Greece. 2021. In *Advances in Longitudinal Data Methods in Applied Economic Research: 2020 International Conference on Applied Economics (ICOAE)*. Cham: Springer International Publishing.
Petrakos, George, Konstantinos Rontos, Chara Vavoura, and Ioannis Vavouras. 2022. Quantifying the Effects of recent Economic and Fiscal Crises on Income Inequality in Greece. Paper presented at 2022 International Conference on Applied Economics (ICOAE), Madrid, Spain, July 7–9.
Petrakos, George, Konstantinos Rontos, Chara Vavoura, Luca Salvati, and Ioannis Vavouras. 2021b. Political budget cycles and the effects of the excessive deficit procedure: The case of Greece. *Regional Statistics* 11: 32–51. [CrossRef]
Petrakos, George, Konstantinos Rontos, Chara Vavoura, Luca Salvati, and Ioannis Vavouras. 2023. Income inequality in the over-indebted Eurozone countries and the role of the Excessive Deficit Procedure. *Open Economies Review*, *forthcoming*.
Piketty, Thomas, and Gabriel Zucman. 2014. Capital is back: Wealth-income ratios in rich countries 1700–2010. *The Quarterly Journal of Economics* 129: 1255–310. [CrossRef]
Rogoff, Kenneth, and Anne Sibert. 1988. Elections and macro-economic policy cycles. *Review of Economic Studies* 55: 1–16. [CrossRef]
Sever, Can, and Emekcan Yucel. 2021. *Electoral Cycles in Inequality*. Working Papers 2021/01. Istanbul: Bogazici University, Department of Economics.
Tsakloglou, Panos. 1990. Aspects of poverty in Greece. *Review of Income and Wealth* 36: 381–402. [CrossRef]
Tsakloglou, Panos. 1993. Aspects of inequality in Greece: Measurement, decomposition and inter-temporal change. *Journal of Development Economics* 40: 53–74. [CrossRef]
Tsakloglou, Panos, and Giota Panopoulou. 1998. Who are the poor in Greece? Analysing poverty under alternative concepts of resources and equivalence scales. *Journal of European Social Policy* 8: 213–36.

Tsakloglou, Panos, and Theodore Mitrakos. 2006. Inequality and poverty in Greece in the last quarter of the twentieth century. In *Social Policy in Greece*. Edited by Maria Petmesidou and Elias Mossialos. Aldershot: Ashgate, pp. 126–43.

Tsiotas, Dimitrios, Thomas Krabokoukis, and Serafeim Polyzos. 2020. Detecting interregional patterns in tourism seasonality of Greece: A principal components analysis approach. *Regional Science Inquiry* 12: 91–112.

Vavoura, Chara, and Ioannis Vavouras. 2022. Income inequality and poverty in Greece during the recent economic, fiscal and COVID-19 crises. *Social Cohesion and Development* 17: 5–21.

Zandi, GholamReza, Raja Rehan, Qazi Muhammad Adnan Hye, Sadaf Mubeen, and Shujaat Abbas. 2022. Do Corruption, Inflation and Unemployment Influence the Income Inequality of Developing Asian Countries? *International Journal of Applied Economics, Finance and Accounting* 14: 118–28. [CrossRef]

Zeqiraj, Veton, Kazi Sohag, and Shawkat Hammoudeh. 2022. Financial inclusion in developing countries: Do quality institutions matter? *Journal of International Financial Markets, Institutions and Money* 81: 101677. [CrossRef]

MDPI AG
Grosspeteranlage 5
4052 Basel
Switzerland
Tel.: +41 61 683 77 34

Economies Editorial Office
E-mail: economies@mdpi.com
www.mdpi.com/journal/economies

Academic Open
Access Publishing

mdpi.com

ISBN 978-3-7258-2442-7