PhD Course in
Labor, Development and Innovation

PhD School in
Engineering For Economics – Economics for Engineering

XXXVII cycle

University of Modena and Reggio Emilia

Granular Policy Evaluation: Socio-Economic and Spatial Analysis in an Unequal Country

---

PhD Candidate
Gianluca Monturano

Tutor
Professor Giovanni Solinas

Co-Tutor
Professor Fabrizio Patriarca

PhD Coordinator
Professor Ylenia Curzi

"E se l'interesse nell'appassionarsi allo studio
dell'economia consistesse nella speranza
che la povertà e l'ignoranza possano
essere gradualmente eliminate?"
- Federico Caffè

Acknowledgements

First, I would like to thank my PhD Tutors, Professor Giovanni Solinas and Professor Fabrizio Patriarca, for their precious and constant support throughout my PhD. Their expert suggestions and timely advice have been an essential guide for my research work and for my academic and personal growth. Their attention and availability have accompanied me in every phase of this experience, contributing significantly to the achievement of my scientific goals. A special thanks goes to the Department of Economics of the University of Modena and Reggio Emilia, to the "Marco Biagi" Foundation and to all the Professors, researchers and staff that I have had the pleasure of meeting during these years. Their support, together with the climate of collaboration and discussion that they have been able to create, has been fundamental for my academic development. A special thought also goes to my PhD colleagues, who have made this experience unrepeatable. A sincere and profound thanks goes to Professor Giuliano Resce. It is thanks to his trust in my abilities that I have managed to reach goals that at the beginning seemed inaccessible. I also thank him for sharing opportunities, knowledge, ideas and collaborations with me, greatly enriching my academic and human background. I also express my enormous gratitude to my co-authors and colleagues, now also great friends, and more, including Drs. Lorenzo Di Stefano, Valentina Erasmo, Giulia Valeria Sonzogno, Professors Amedeo Argentiero, Giovanni Cerulli, Giuseppe Coco, Patrizio Frederic, Giovanni Gallo, Paolo Maranzano, Luigi Mastronardi, Giulio Pedrini, Pasquale Tridico and Marco Ventura, as well as to all the members of the GRAPE group. The work shared with them and the continuous exchange of ideas have been an experience of inestimable value, both on a professional and human level. I would like to thank Professor Michele Cincera for offering me the opportunity to carry out the Visiting Research in Brussels at the Université libre de Bruxelles. A very special thanks goes to Professor Angela Stefania Bergantino and to all my colleagues at the Department of Economics, Management and Business Law of the University of Bari, for the trust they have placed in me, even before the conclusion of the PhD. Last but not least, I would like to thank my family. In particular, my parents, my brother, my grandparents and my uncles, for their unconditional love and their continuous support in every moment of my life. Their presence, their encouragement and their trust have been the strength that has allowed me to face with determination every step of this journey. A special thanks also goes to my friends, who with their affection and closeness have made this trip unforgettable.

Contents

List of Figures

List of Tables

List of Equations

Abstract

This doctoral thesis focuses on the evaluation of public policies at the municipal level in Italy, with specific attention to territorial cohesion policies and income support measures. The first Chapter analyzes the impact of the funds allocated within the framework of the National Strategy for Inner Areas, a place-based policy aimed at combating depopulation in the most remote and marginal municipalities. Using a Staggered Difference-in-Difference econometric model, the study evaluates the effects of financial support on the demographic structure and the number of economic activities at the municipal level. The results show a significant increase in new business units in the treated municipalities, without notable changes in the population structure. Additionally, the presence of spillover effects on neighboring municipalities is analyzed, revealing a positive impact on the increase in economic activities in the surrounding areas. The second Chapter investigates the features of delays in cohesion infrastructure projects in Italy. Using a rich dataset at the project and municipal levels, the analysis predicts the socio-economic and institutional features relevant to project timelines through the use of advanced Machine Learning techniques. The efficiency of educational institutions, administrative capacity, and allocated funds are found to be key factors that significantly influence delays. The study further demonstrates that Machine Learning can improve the efficiency of public investments by overcoming logistical and administrative barriers associated with the implementation of cohesion policies. The third Chapter examines the impact of poverty and inequality on the distribution of the Citizenship Income at the municipal level, a minimum income policy for families in economic hardship. Using spatial econometric models, the research explores the correlation between poverty, inequality, and program participation. The results indicate that areas with greater socio-economic difficulties show higher participation, while wealthier regions exhibit lower participation. Furthermore, it is noted that the COVID-19 pandemic has exacerbated territorial disparities, making these dynamics even more pronounced. The fourth Chapter explores the relationship between the Citizenship Income and voting behavior, taking into account the socio-spatial differences across Italy. The study combines administrative data on program beneficiaries with municipal-level electoral data. Using a continuous treatment Difference-in-Difference model, the analysis investigates how financial support has influenced voting for the governing party. The results show that in disadvantaged areas with lower institutional quality, the policy has increased electoral support for the governing party, while in more prosperous areas, the opposite effect was observed. In conclusion, the thesis provides an in-depth analysis of public policies. The use of econometric, counterfactual, and spatial models, along with predictive Machine Learning algorithms, supported by granular municipal data, has enabled a detailed examination of socio-economic disparities in Italy, surpassing the traditional analyses based on regional and macro-regional divisions. The research highlights how cohesion and income support policies can be optimized through targeted planning tailored to local specificities. The findings offer useful policy insights for policymakers, providing concrete tools to reduce territorial inequalities and promote more balanced and inclusive development across the country.

Introduction and Conclusion

Introduction to Spatial Inequalities and Framework of the Thesis

In recent decades, the issue of socioeconomic and territorial inequalities has emerged as one of the most significant challenges for policy makers around the world. In particular, with the globalisation of markets and technological growth, the differences between different social groups and regions of the world have become more pronounced, between those who are able to obtain benefits from these processes (winners) and those who are excluded (losers) (Milanovic, 2016; Acemoglu and Restrepo, 2018). In this sense, globalisation has certainly facilitated an intense interconnection between economies at a global level, but it has also given rise to a concentration of wealth that is not spatially uniform between urban areas, which are more integrated and agglomerated, and peripheral regions, which remain predominantly cut off from these flows of progress and wealth. This polarisation is, therefore, not limited to separating rich countries from poor ones, but penetrates deeply within individual nations, determining what economists define as internal inequality within countries (Milanovic, 2016). At a general level, economic elites and already developed regions have tried to maximise and exploit to the maximum the advantages offered by globalisation, while less developed areas have seen their conditions of marginality and peripherality worsen (Rodríguez-Pose, 2018). This evidence has been widely demonstrated in economic literature. In particular, Piketty (2014) in his fundamental work "Capital in the twenty-first century" illustrates how, in different economic contexts, the return on capital tends to regularly exceed economic growth rates, determining a concentration of unequal wealth, increasingly in the "hands of a few" (Krugman, 1991), in his seminal work on the new theory of economic geography and trade, has highlighted how, at the spatial level, agglomeration in many urban areas leads to virtuous cycles of growth, which attract further investment and human capital, while less developed areas tend to enter a vicious circle of stagnation and long-term decline. This imbalance not only limits economic development but also undermines social cohesion, creating scenarios of 'divided nations', where territorial disparities and social exclusion are marked and persistent (Rodríguez-Pose, 2018; Rodríguez-Pose et al., 2021). Stiglitz (2012) has extensively discussed the consequences of inequalities in the socio-economic system, arguing that in the absence of effective economic policies, markets tend to increase inequality, rather than reduce it. These analyses support the idea that redistributive policies and targeted investments are essential to counter growing inequalities. These theses have also been confirmed by Atkinson and Bourguignon (2014) who, in a large work focused on the study of income distribution, argue that the growing concentration of wealth requires targeted redistributive policies, capable of ensuring equity and social cohesion. Milanovic (2016), analysing the changes in income structure in recent decades, discusses the factors that have contributed to these changes, arguing that economic policies and regulations, at the international level, must be revised to address the growth of income inequalities, for example, through a new global taxation procedure, more flexible migration policies, increases in international aid, investments in services such as education and health, and the promotion of inclusive institutions. In this sense, in a broad line of works Acemoglu et al. (2001, 2003, 2004) and Acemoglu and Robinson (2008, 2013), exploring the role of institutions, distinguish between "extractive" institutions, which concentrate power and wealth, and "inclusive" ones, which instead, promote equity and innovation. They suggest that to effectively address inequalities, it is now necessary to promote institutions that incentivise participation and equality of opportunity.

In general, inequalities represent, not only an economic challenge, but also political and social issues. In fact, unequal societies tend to be less cohesive, showing greater social tensions and less political stability. This is particularly true in periods of increasing polarisation, where economic disparities can easily translate into broader political/social conflicts (Rodríguez-Pose et al., 2021).

Effectively addressing inequalities, therefore, requires an organised approach which takes into account the multidimensionality of the problem. The discussion on how to structure these policies is crucial and necessitates a continuous dialogue between economic theory, empirical evidence, and political practice. From this point of view, many institutions have launched policy programmes to try to mitigate the effects of inequality. Stimulated by the Barca (2009) Report, the European Union endeavoured to focus its economic and territorial cohesion policies, aimed at local development, by assigning importance to interventions sensitive to the territorial context and the endogenous potential of individual territories, suggesting that "place-based" interventions, or policies based on the specificities of places, can be more effective in responding to the needs of disadvantaged areas, so as to promote a more homogeneous balance in the various national contexts (Barca et al., 2012).

At a spatial level, our country, Italy, appears to be, among the advanced nations, one of the countries with the greatest territorial disparities. In fact, in Italy, socioeconomic inequalities represent one of the most complex challenges to be faced.

The dualism between the regions of the North, traditionally more industrialised and economically advanced, and those of the South, backward and less developed, reflect a framework of inequality that has deep historical roots, which continues to influence national development policies (Lagravinese, 2015; Felice, 2019). This territorial dichotomy is not only a question of economic differences, but also extends to social, infrastructural, and cultural aspects, contributing to a complex network of interdependencies, which influence access to opportunities and the quality of life of local populations. However, today, despite these long-term macro-regional disparities, the persistence of the economic crisis has revealed the need to deepen the study of territorial disparities in our country also, overcoming the historical dualism between the North and the South. Indeed, there are marked differences between Italian urban areas, which are more integrated and agglomerated and which benefit from privileged access to economic, social and service networks, and peripheral areas, which, although rich in cultural and landscape potential, are largely underused and continue to lag behind in terms of development and growth. In particular, these marginal areas tend to suffer from geographical isolation and lack of access to essential services. In terms of local development, the growing gap between areas has, not only economic repercussions, but also risks fuelling a vicious circle of demographic decline and depopulation, especially in smaller communities.

The lack of development policies that take into account local specificities risks undermining the very cohesion of the Nation, creating a two-speed Italy, where peripheral areas suffer a slow decline. In response to these issues, following Barca et al. (2014), in 2014, the Italian government approved a specific place-based policy: the National Strategy for Inner Areas (SNAI), which divides the Italian territory between agglomerated urban centres capable of providing essential services, and marginal inner areas, isolated in terms of accessibility and in socio-economic and demographic decline. In terms of policy, the SNAI represents a significant innovation within the panorama of Italian policies, since, from a bottom-up perspective, it changes the paradigm of local development policies by moving from classical centralised policies that, as Iammarino et al. (2017) observe, often tend to ignore local specificities, ending up favouring mainly already developed areas. In contrast, the SNAI model aims to create development conditions that can make the Inner Areas economically and socially sustainable.

The adoption of a cohesion policy such as the SNAI not only represents a concrete response to the need to promote the economic development of marginal areas, but also offers an interesting laboratory to evaluate the effectiveness of alternative approaches to traditional administrative centralism. In fact, the SNAI allows us to evaluate whether a territorially localised policy approach can stimulate self-sufficient growth, while reducing the dependence of these areas on subsidies and external transfers, which often produce limited benefits in the long term.

To date, studies conducted on the SNAI have mainly been concerned with the use of municipal classification for the study of income, inter-municipal cooperation, and population structure. At a political-administrative level, the strategy presents, as in general do many cohesion policies, delays and problems in implementation and, to date, its implementation is much debated. In the light of this, the first Chapter (1) of this thesis focuses on an in-depth counterfactual econometric analysis of the effectiveness of the SNAI, assessing the extent to which the programme objectives have been achieved and its impact on key indicators such as population density, economic development rate, and perception of quality of life in the beneficiary municipalities and surrounding areas. This analysis will provide an empirical basis for understanding whether and how place-based policies can contribute to solving Italy's territorial challenges.

A theme which is closely linked to the reduction of territorial disparities, i.e. the improvement of social cohesion between areas and the socioeconomic conditions of local communities, is that of infrastructure, which represents one of the main levers for ensuring mobility, attracting investments and, consequently, promoting access to essential services.

The lack of basic infrastructure, such as transport networks, digital connections, and health services, discourages private investments, and limits the mobility of people, contributing to the progressive deterioration of public services and, therefore, to the phenomenon of depopulation of local communities.

The European Union's cohesion policy allocates a significant portion of its budget - approximately one-third of the total - to spending to reduce the gaps between European regions. These resources, to a large extent, finance specific infrastructure interventions. In terms of implementing infrastructure cohesion projects, Italy faces significant challenges, revealing significant delays and inefficiencies in the management of public resources. Delays in the construction of infrastructure not only slow down the economic development of disadvantaged areas but also generate a significant economic and social cost for the State. As highlighted by Cerqua and Pellegrini (2017), administrative inefficiency and the complexity of bureaucratic procedures are among the main causes of infrastructure delays, since they compromise the quality of life in less developed areas and reduce the competitiveness of the entire country.

Despite a rich literature questioning the effectiveness of European cohesion funds, to date, there are not many research works that study the delays of infrastructure projects, especially in predictive terms. Given the importance of the topic, in addition to the new spending programmes in progress, the second Chapter (2) of the thesis intends to systematically address the problem of infrastructure delays, exploring in detail the factors that influence their timing and efficiency. Using the latest predictive Machine Learning (ML) algorithms, the analysis aims to identify the key variables that determine delays, proposing possible solutions to improving the management and effectiveness of public resources.

In addition to the territorial inequalities highlighted, a closely related and equally pressing problem is poverty, which continues to be one of the main challenges for modern societies, especially in economically backward contexts or in areas particularly affected by economic crises (Saraceno et al., 2020). Poverty not only limits access to essential resources such as food, housing, and medical care, but also compromises opportunities for personal and social development for entire communities, in a cycle of deprivation that is transmitted at an intergenerational level.

In 2019, Italy equipped itself with a national guaranteed minimum income scheme: the Italian Citizenship Income. The approval of this policy marked an important step in the Italian strategy to combat poverty, inequalities, and social exclusion, since Italy was one of the last countries to equip itself with protection tools to combat these phenomena. The Italian Citizenship Income, financed with significant national resources (more than 6 billion Euros), was conceived not only as a measure of economic support, but also as a work reintegration program, with the aim of reducing the multidimensional problems present in the country and offering new opportunities to disadvantaged families. During its period of operation, until December 2023, the policy was the subject of numerous criticisms and discussions. In particular, there was much debate regarding the group of beneficiaries who participated in the programme and, therefore, had obtained the subsidy, which highlighted a very unbalanced distribution towards the regions of Southern Italy. In these terms, the welfare literature, for example Esping-Andersen (1990), suggests that the effectiveness of support policies depends greatly on the ability to adapt interventions to the specific characteristics of territorial areas, rather than being based on a uniform approach, and raises the question of whether and how welfare programmes can be improved to better address the needs of the most fragile territories. In this line of studies, the third Chapter (3) of the thesis evaluates, at a spatial level, the influence of poverty and inequalities at the municipal level in obtaining the subsidy. Using spatial regression models, we identify how territorial disparities produce improvements in political participation.

Also, in relation to the discussions that revolve around the implementation of the Citizen's Income, the policy has been much debated regarding potential electoral effects. A large number of political economy studies, see for example Lewis-Beck (1985) and Lewis-Beck and Stegmaier (2018), show that the provision of economic subsidies often has an influence on the electoral preferences of the population, helping to build political consensus for the forces promoting welfare measures.

In Italy, public opinion has long focused on the study of possible positive effects of the Citizenship Income on political consensus for the incumbent party of the measure, especially in the most disadvantaged areas of the country, where it has been mostly disbursed. This phenomenon, studied in the fourth chapter (4), raises important questions on how welfare can influence electoral dynamics and offers insights into understanding the potential of income support policies as a tool for electoral advantage, particularly in areas where the impact of welfare is more tangible.

Objectives and Methodology of the Thesis

This thesis aims to examine the effectiveness of Italian public policies in promoting territorial cohesion and socioeconomic support, with a rigorous and quantitative methodological approach, employing counterfactual econometric models, machine learning algorithms, and spatial regressions. This ensures robust and reliable results for a fine-grained understanding of the socioeconomic dynamics and policy challenges characterising Italy. The specific objectives of the thesis work include:

1. Short-term evaluation of the National Strategy for Internal Areas (SNAI);
2. Determinants of delays in infrastructure cohesion projects;
3. Influence of inequalities and poverty at local level on participation in the Citizenship Income;
4. Impact of the Citizenship Income on electoral behavior and political consensus of the proposing list.

This thesis makes an original contribution to the literature on cohesion and welfare policies adapted to the territorial context, demonstrating how targeted interventions can reduce inequalities and promote sustainable development. The results obtained offer concrete insights for future policy decisions, underlining the effectiveness of an approach that combines economic and infrastructural support in response to territorial specificities.

Conclusion

This thesis has explored, in depth, the impact and effectiveness of Italian public policies focused on territorial cohesion and socioeconomic support, with a particular focus on the National Strategy for Internal Areas (SNAI), delays in infrastructure cohesion projects, and the influence of welfare policies, such as the Citizenship Income. Using advanced quantitative methodologies, including spatial and counterfactual econometric analyses, as well as Machine Learning (ML) algorithms, this research has provided a detailed examination of the current socioeconomic dynamics and policy challenges in Italy, offering crucial insights for targeted policy interventions and improvements in policy strategies.

In particular, one of the distinctive features of this thesis is the use of granular data at municipal and project levels, a significant advance compared to previous works conducted at more aggregated territorial scales. This approach allowed for a more detailed understanding of local dynamics and provided a contextually relevant perspective on how public policies directly impact communities at the territorial level.

Specifically, the SNAI study highlighted significant improvements in the availability of services and the stimulation of economic growth in internal areas. Despite some implementation challenges, the municipalities involved recorded positive impacts emanating from the policy treatments with increases in economic activity. The chapter on delays in infrastructural cohesion projects highlighted administrative inefficiencies of educational institutions and problems related to the quality of institutions as the main obstacles.

The analysis conducted with ML algorithms identified the main causes of these delays and suggested that improvements in project management and greater transparency could accelerate the implementation of the infrastructure needed for balanced territorial development.

With regard to the Citizenship Income, a positive relationship was revealed between areas with high levels of poverty and low levels of income with respect to participation in the programme. By contrast, in areas with greater economic inequalities, non-homogeneous effects on a territorial scale are found, connected to the very nature of the policy which was conditioned on compliance with specific income requirements. These results raise important questions concerning the adaptability of welfare policies to individual local specificities.

Furthermore, the political influence of the Citizens' Income was examined, highlighting that this measure is able to modify electoral preferences, benefiting, in terms of consensus, the list proposing the measure, especially in territorial areas that suffer the weight of long-term socioeconomic problems. This suggests that welfare can function as a political instrument that is potentially able to influence voting tendencies, strengthening the interconnection between economic support and electoral behaviour.

The recommendations and policy ideas that emerged from this study include the need for a profound review of current policies to identify areas of strength and weakness. It is crucial to improve administrative management and reduce bureaucratic complexity in infrastructure projects. Furthermore, for welfare measures, it is recommended to better consider local conditions in implementation strategies to ensure that benefits are fairly distributed.

In conclusion, this thesis provides a significant contribution to the understanding of cohesion and welfare policies in Italy, highlighting persistent challenges and outlining opportunities for the future. The research highlighted the need for a continuous updating of policy strategies to effectively address socioeconomic inequalities, supporting the importance of an innovative and well-coordinated approach that integrates economic theory, empirical evidence, and policy practice to promote more equitable and inclusive development, and also in consideration of the new policy lines and programming that have been activated and that will be activated in the near future.

Thesis Structure

The thesis is divided into four Chapters and is organised as follows. Chapter 1 focuses on the analysis of the impact of the National Strategy of Inner Areas, in its first years of activity. In particular, the effects of the SNAI in terms of depopulation and dynamics of economic activities in the treated municipalities and in neighbouring municipalities are examined, using a difference-in-differences econometric model with staggered treatment. The results confirm that the SNAI has generated positive effects in terms of economic activities in the treated municipalities, as well as benefits for municipalities not directly involved, suggesting a positive spillover impact. Chapter 2, through the use of predictive Machine Learning algorithms, identifies the critical factors that influence delays in infrastructure projects, with reference to the 2007-2013 territorial cohesion cycle. This in-depth analysis, at project level, represents a resource to determine the socioeconomic characteristics that can optimise administrative processes and improve the management of public infrastructures, helping to reduce costs and implementation times. Chapter 3 explores, at the municipal level, how socioeconomic conditions and poverty influence the distribution of the Citizens' Income, through spatially clustered regression models. The results indicate that the poorest areas benefit more from the policy, highlighting the need for targeted policies to reduce territorial disparities. Chapter 4 analyses how the Citizen's Income influences electoral preferences at the local level, demonstrating that the policy had a positive impact on the political consensus of the proposing list, especially in Southern Italy.

Bibliography

Acemoglu, D., Johnson, S., and Robinson, J. (2004). Institutions, volatility, and crises. *Growth and productivity in East Asia*, 13.

Acemoglu, D., Johnson, S., Robinson, J., and Thaicharoen, Y. (2003). Institutional causes, macroeconomic symptoms: volatility, crises and growth. *Journal of monetary economics*, 50(1):49–123.

Acemoglu, D., Johnson, S., and Robinson, J. A. (2001). The colonial origins of comparative development: An empirical investigation. *American economic review*, 91(5):1369–1401.

Acemoglu, D. and Restrepo, P. (2018). The race between man and machine: Implications of technology for growth, factor shares, and employment. *American economic review*, 108(6):1488–1542.

Acemoglu, D. and Robinson, J. A. (2008). Persistence of power, elites, and institutions. *American Economic Review*, 98(1):267–293.

Acemoglu, D. and Robinson, J. A. (2013). *Why nations fail: The origins of power, prosperity, and poverty.* Crown Currency.

Atkinson, A. B. and Bourguignon, F. (2014). *Handbook of income distribution.* Elsevier.

Barca, F. (2009). An agenda for a reformed cohesion policy. a place-based approach to meeting european union challenges and expectations. Technical report, European Commission. Bruxelles.

Barca, F., Casavola, P., and Lucatelli, S. (2014). Strategia nazionale per le aree interne: definizione, obiettivi, strumenti e governance. *Materiali ULAV*, 31(10).

Barca, F., McCann, P., and Rodríguez-Pose, A. (2012). The case for regional development intervention: place-based versus place-neutral approaches. *Journal of regional science*, 52(1):134–152.

Cerqua, A. and Pellegrini, G. (2017). Industrial policy evaluation in the presence of spillovers. *Small Business Economics*, 49:671–686.

Esping-Andersen, G. (1990). The three worlds of welfare capitalism. *Polity.*

Felice, E. (2019). The roots of a dual equilibrium: Gdp, productivity, and structural change in the italian regions in the long run (1871–2011). *European Review of Economic History*, 23(4):499–528.

Iammarino, S., Rodríguez-Pose, A., and Storper, M. (2017). Why regional development matters for europe's economic future. *European Commission Directorate General for Regional and Urban Policy Working Paper*, 7:2017.

Krugman, P. (1991). Increasing returns and economic geography. *Journal of political economy*, 99(3):483–499.

Lagravinese, R. (2015). Economic crisis and rising gaps north–south: evidence from the italian regions. *Cambridge journal of regions, economy and society*, 8(2):331–342.

Lewis-Beck, M. S. (1985). Pocketbook voting in us national election studies: fact or artifact? *American Journal of Political Science*, pages 348–356.

Lewis-Beck, M. S. and Stegmaier, M. (2018). Economic voting. *The Oxford handbook of public choice*, 1:247–265.

Milanovic, B. (2016). *Global inequality: A new approach for the age of globalization*. Harvard University Press.

Piketty, T. (2014). Capital in the twenty-first century. *Trans. Arthur Goldhammer/Belknap*.

Rodríguez-Pose, A. (2018). The revenge of the places that don't matter (and what to do about it). *Cambridge journal of regions, economy and society*, 11(1):189–209.

Rodríguez-Pose, A., Lee, N., and Lipp, C. (2021). Golfing with trump. social capital, decline, inequality, and the rise of populism in the us. *Cambridge Journal of Regions, Economy and Society*, 14(3):457–481.

Saraceno, C., Benassi, D., and Morlicchio, E. (2020). *Poverty in Italy: Features and drivers in a European perspective*. Policy Press.

Stiglitz, J. E. (2012). *The price of inequality: How today's divided society endangers our future*. WW Norton & Company.

Chapter 1

Revitalizing Italy's Inner Areas

Socio-Economic Impacts of Financial Support and Spatial Spillovers Effects*

Abstract

This study investigates the effects of financial support to Inner Areas leveraging a specific governmental place-based policy to fight depopulation implemented in Italy. Taking advantage of the most recent developments in the econometrics of policy evaluation, we apply a staggered difference-in-difference estimator (DiD) to evaluate the impact of public policy in terms of population structure and the number of plants at the municipal level. The analysis is made possible thanks to a detailed panel dataset containing information about the Italian municipalities over the years 2014-2020. The results show that, over the first two years, the financial support to Inner Areas does not affect the population structure, but it can generate a significant number of extra plants in the treated municipalities. A further key issue is whether the policy can generate spillover effects on neighbours which may either corroborate the encouraging result or invalidate it. To answer this question we combine the baseline model with a spatial empirical strategy, and we find positive spillover effects for extra plants on neighbouring municipalities.

JEL classifications: C21; O12; O18

Keywords: Policy Evaluation; Place-based Policy; Causal Inference; Staggered DiD; Spatial Spillover Effects

---

*This work is a paper co-authored with **Professor Giuliano Resce** (Department of Economics, Via Francesco De Sanctis, 86100 Campobasso CB, and International Inequalities Institute, London School of Economics and Political Science, Houghton Street, London, WC2A 2AE, UK) and **Professor Marco Ventura** (Department of Economic and Law, Sapienza University of Rome, Piazzale Aldo Moro 5, 00185 Roma) and is published in a similar and small version as a Working Document at the Impact Assessment Office, Senate of the Republic, Evaluation Document no. 13, available at https://www.senato.it/japp/bgt/showdoc/19/DOSSIER/0/1382388/index.html?part=dossier_dossier1-sezione_sezione1.

1.1 Introduction

The issue of depopulation in marginal areas has become a critical challenge for policymakers, especially in countries with significant regional disparities such as Italy (Cerqueti and Ausloos, 2015; Felice, 2018; Checchi et al., 2024). This challenge stems from a long history of urban growth juxtaposed with rural underdevelopment, driven by a combination of economic factors such as labour productivity, and the effects of sharing, matching, and learning (Robinson, 1976; Knight, 1976; Fields, 1979; Combes and Gobillon, 2015; Iammarino et al., 2019). These dynamics, along with the technological spillovers, labour pooling, and intermediate input linkages identified by Marshall (1890), have contributed to widening the gap between core and peripheral regions in many countries (Krugman, 1991). In this context, Italy's Inner Areas — marked by geographical isolation and limited access to essential services — have faced population decline and economic stagnation (Gallo and Pagliacci, 2019).[1] Spatial heterogeneity, a major driver of regional income inequality (Bourguignon and Morrisson, 1998), is negatively linked to economic growth (Clarke, 1995; Alesina and Rodrik, 1994; Persson and Tabellini, 1994), making these areas a matter of urgent national concern.[2] In response, the Italian government has introduced place-based policies aimed at revitalising these regions, focusing on financial support and local economic investment. However, the effectiveness of such interventions — particularly in influencing local population dynamics and economic activity — remains an open question. The literature indicates that, despite significant investments, place-based policies in Italy have generally failed to generate substantial economic growth and have sometimes produced unintended negative side effects, such as disproportionately benefitting wealthier individuals (Barone and De Blasio, 2023a,b). This raises important questions about how to design these policies more effectively to ensure equitable benefits and to avoid inefficiencies and adverse outcomes.

This paper contributes to this literature by investigating the short-term effects of initial financial support for Inner Areas. In this task, we leverage a specific place-based bottom-up policy implemented in Italy, called the "National Strategy for Inner Areas" (*Strategia Nazionale Aree Interne* - SNAI) (Barca et al., 2014). The policy was introduced in 2014 but, due to administrative, bureaucratic, and financial problems, it commenced

[1] The concept of Inner Area was first introduced into the Italian discourse in the late 1950s, primarily in the context of the *Mezzogiorno's* challenges. During this period, Manlio Rossi Doria, an agrarian economist, coined the term Inner Area to denote rural regions characterised by low productivity, often resulting from an inadequate irrigation system. While initially associated with the conditions of Southern Italy, it became evident that the process of marginalisation affecting these Inner Areas extended to other parts of the country. Rossi Doria pointed out that similar issues were prevalent in mountainous and hilly regions of the Alps in the North and the Apennines in the Centre of Italy.

[2] Recent studies have underscored the conception that the rise in within-country inequality, which has coincided with a reduction in cross-country inequality in recent decades (Arestis et al., 2011; Liberati, 2015), may be a significant driver behind the global populist backlash (Goodhart, 2017; Piketty et al., 2018; Rodríguez-Pose, 2018; Albanese et al., 2022).

operating, i.e. financing municipalities, only as late as 2018. In addition to this delay, the new European programming 2021-2027 has brought changes to the policy, with a structural review still ongoing. The central aim of the policy is to combat depopulation. Consequently, it is rather unlikely that significant effects will be seen in the short term. However, the absence of positive effects on depopulation in the short term does not necessarily imply a general failure of the policy. Moreover, given the amount of money at stake (over 1 billion euros of national funds and European structural funds) it is still worth evaluating the policy in terms of "clues", i.e. indicators that can tell us something about the possible future success or failure of the policy. As a first test, we focus on the population structure using net immigration flows per 1,000 inhabitants (the municipal-level number of inflows minus the municipal-level number of outflows divided by population). We selected net immigration flows to assess whether the new economic opportunities, linked to the SNAI policy, produced any stimulus to new immigration flows, which could have the potential for refraining from the demographic decline characterising Inner Areas (Basile et al., 2023). As a second test, we analyse the policy in terms of the number of business sites of firms that, consistent with Criscuolo et al. (2019) and Gibbons et al. (2019), we refer to as "plants". We have chosen such an outcome to evaluate the effects of the policy, as it is one of the most sensitive to institutional changes in the short run (Lu et al., 2019). Effects on the number of plants can be, by far, more immediate than those on population, i.e. detectable in the short run. However, we stress that an increase in the number of plants is only a necessary but, in itself, an insufficient condition for the success of the policy. Indeed, finding positive effects on the number of plants in the short run generates a reasonable expectation of finding positive effects on the population in the long run. By the same token, if no effects are found on plants in the short run, it is quite unlikely that depopulation can be restrained in the long run.

Inner areas are territories characterised by being located at a significant distance from the main centres providing essential services to inhabitants, i.e., health, education, and mobility. In this regard, Inner Areas are defined at the municipal level according to a distance criterion: a municipality is classified as an inner area if it is more than 20 minutes away from secondary schools, hospitals, and train stations. Consistently with the policy design, our analysis is carried out at the municipality level, the most disaggregated administrative level in Italy (Barca et al., 2014). Previous studies have analysed the impacts of regional policies at aggregated levels (Rodríguez-Pose and Fratesi, 2004; Becker et al., 2010; Mohl and Hagen, 2010), particularly by region (i.e., NUTS2), but currently, the increasing availability of detailed municipality-level data allows us to disentangle direct impacts of the policies that otherwise would not be observable at finer levels of aggregation. These fine pieces of information, in turn, allow for a fine-tuning of the policies that greatly benefit the reliability of the analysis.

We use the data made publicly available by the Ministry for Territorial Cohesion, described in the "Annual reports on the implementation of the policy"[3] which identifies the municipalities that have received SNAI funding (that, following our econometric empirical design, illustrated in Section 1.5, we define as 'treated') and the year of first entry to the first funding, for the years 2018, 2019 and 2020 (Dipartimento per le politiche di coesione, 2019, 2020, 2021). As detailed in Section 1.3, the policy has been implemented in a multi-stage process and a limited number of municipalities have been financed up to 2020. By construction, these municipalities are similar in terms of observable characteristics and differ only by the treatment timing. From a methodological point of view, this staggered entry allows us to take advantage of one of the most recent developments in the econometrics of policy evaluation. Notably, we use the estimator developed by Sun and Abraham (2021) which generalises the difference-in-difference (DiD) estimator accounting for staggered entry into treatment. Other similar staggered difference-in-difference estimators are also used, such as Gardner (2022), Callaway and Sant'Anna (2021), and Borusyak et al. (2021), to check that the results obtained still apply.

We find no significant effects of the financial support in terms of population structure, i.e., the net immigration flows. The absence of effects can be partially due to the short period of analysis which does not allow us to capture long-run effects of the policy. Focusing on extra plants, the results show that, on average, the policy has produced results since its inception, generating a positive effect in the treated municipalities over the first two years. However, this result, in turn, raises another crucial question about the presence of spillover effects. Positive spillovers generated on neighbouring municipalities would reinforce the effectiveness of the policy, but negative spillovers would bias the estimates upwards raising doubts about its actual effectiveness. To investigate this important issue in detail, we follow the empirical approach proposed by Kline and Moretti (2014a) finding positive spillover effects. This result corroborates the validity of extra plants generated by the policy and is in line with Kline and Moretti (2014a).

The remainder of the paper is structured as follows: the next Section introduces a literature review on the effects of place-based policies; Section 1.3 presents the institutional setting; Section 1.4 summarises the dataset; Section 1.5 describes the empirical strategy; Section 1.6 shows the results and Section 1.7 concludes the study.

1.2 Literature Review on Place-Based Policies

The growth of inequality between territorial areas and countries (Rodríguez-Pose and Hardy, 2015; Iammarino et al., 2019), assisted by globalisation and technical progress (for further information see for example Acemoglu (2002), Capello et al. (2011), Paba and Solinas (2018), Acemoglu and Restrepo (2020)), has, for some time, pushed many national and international institutions, such as the European Union (EU), to develop consistent cohesion

[3]The latest available Annual Report refers to 2020.

policy frameworks. These policies have, as their main objective, the reduction of economic disparities, generally resulting from geographical remoteness (Becker et al., 2010; Farole et al., 2011; Cerqua and Pellegrini, 2014; Albanese et al., 2023). They are implemented using specific centralised instruments that, invariably, have attracted both supporters and opponents over time. On the one hand, a cohesion policy has been recognised as necessary to compensate the most backward regions for the negative effects that the reduction in barriers has had on their economies. On the other hand, it has been considered to have been a profligate waste of resources, with high costs in terms of efficiency and, consequently, economic growth.

Currently, cohesion measures are increasingly changing their direction towards "place-based policies", which emphasise the specific potential of each area. Therefore, they require interventions adapted to local, natural, and institutional resources. They are structured with horizontal cooperative governance between all the economic and social actors involved (Barca, 2009; Farole et al., 2011; Barca et al., 2012; Neumark and Simpson, 2015).

A significant body of scientific work, such as Farole et al. (2011), Barca et al. (2012), and Neumark and Simpson (2015), considers those policies capable of addressing delays in regional development, especially in the case of marked spatial heterogeneity. The most widespread place-based policies concern programmes that offer direct incentives for companies and production plants located in economically weak regions and areas. Principal examples of these policies include Urban Empowerment Zones (UEZ), Enterprise Zones (EZ), Special Economic Zones (SEZ), EU Structural Funds (ESF), Industrial Cluster policies, and bottom-up policies for local communities (Barca, 2009; Accetturo and De Blasio, 2012; Kline and Moretti, 2014b; Neumark and Simpson, 2015; Lu et al., 2019).

At an empirical level, great importance is given to the study of the effects and impacts that these policies produce on socio-economic outcomes, see for example Bartik (2020), Moretti (2024), such as employment (Austin et al., 2018), wages (Busso et al., 2013), entrepreneurship, with the creation of new businesses or production plants and the strengthening of existing economic activities (Accetturo and De Blasio, 2012; Criscuolo et al., 2019; Lu et al., 2019), and on other economic development phenomena, such as depopulation and house prices (Busso et al., 2013; Ciani and De Blasio, 2015; Koster and Van Ommeren, 2019). Furthermore, scientific relevance is also attributed to the study of any spatial spillover effects deriving from the policies (Cerqua and Pellegrini, 2017; Zheng et al., 2017). It has been highlighted that economic activities could react to the policies by moving from other regions or territories adjacent to the target areas, thus taking away benefits associated with the program without improving the well-being of local residents (Kline, 2010; Ehrlich and Seidel, 2018; Criscuolo et al., 2019). Similarly, the policy could generate positive effects for untreated neighbours (Greenstone et al., 2010; Kline and Moretti, 2014a).

The scientific literature on place-based policies shows discordant and heterogeneous econometrics results on a spatial level and on different socio-economic variables (Kline and Moretti, 2014a; Cerqua and Pellegrini, 2017; Becker et al., 2018).

Positive results for these policies, although often nuanced and limited, can be observed in several advanced countries and specific territorial contexts. For example, Pellegrini et al. (2013) and Crescenzi and Giua (2020), from a macro perspective, find, albeit to a moderate extent and with significant regional variations, that the European Union Structural Funds support employment and economic growth in the main EU countries. In particular, Pellegrini et al. (2013) assess the impact of European Regional Policy on economic growth from 1994-2006 and show a positive effect of the policies on per capita GDP in the target regions, confirming the effectiveness of regional policy to favour economic growth in the less developed regions. On the other hand, Crescenzi and Giua (2020) analyse the effect of the funds on economic growth at a spatial level. They find general positive impacts of cohesion funds on economic development, with impacts varying significantly on a geographical scale.

In addition to economic growth, other scientific works show that place-based policies have positive and significant impacts, in target regions, also on wellbeing (Blouri and Ehrlich, 2020; Ferrara et al., 2022) and in specific economic and service sectors, such as innovation and transport accessibility (Ferrara et al., 2017).

Criscuolo et al. (2019) investigate the causal effects of subsidies aimed at labour and industrial development in disadvantaged areas of the United Kingdom[4], finding a significant increase in investment and employment, especially for small firms and plants in the target areas. Additionally, this work notes that there are no negative spillover effects in terms of displacement effects of workers from ineligible areas to treated areas, confirming a general positive effect of the policy. Austin et al. (2018) evaluate the effectiveness of place-based pro-employment policies targeted at depressed regions of the United States characterised by high unemployment rates. They observed that the expansion of interventions aimed at increasing labour demand can significantly reduce unemployment and stimulate local economic growth in the target areas. In the Chinese context, encouraging results are recorded, especially with regard to the impact of the Special Economic Zones. Lu et al. (2019), starting from an empirical difference-in-difference estimation design, study the link between SEZs approved in China and local economic growth. They detect significant increases in capital investment and employment, growth in productivity and workers' wages, as well as an increase in the number of firms and plants in the treated areas. These last results are confirmed also by Zheng and Pan (2024), which show that SEZs significantly stimulate the birth of new high-value-added firms in the service sector and, at the same time, favour manufacturing activities. From a spatial perspective, these works highlight that the positive impacts of SEZs depend on their geographical location, the type and extent of the programme, and the industrial sector they target (Lu et al., 2019; Zheng and Pan, 2024).

[4] Regional Selective Assistance (RSA).

At a more granular level, Giua (2017), using Spatial Regression Discontinuity Designs (RDDs), evaluate the impact of the Structural Funds for Objective 1[5], with the intention of stimulating the economic development of the lagging Italian regions. The study shows positive impacts on employment levels in the treated areas of Objective 1, also noting the absence of displacement effects of economic activities from nearby untreated areas. The study also highlights that the positive impacts were concentrated in areas with higher treatment intensity.

Other studies showing positive outcomes of place-based policies in disadvantaged areas have focused on the impacts of specific programmes. For example, Cusimano et al. (2021) analyses the effectiveness of an Italian place-based regional development programme (the "Integrated Territorial Projects"), using a dose-response function, which captures treatment heterogeneity in terms of intensity. The programme is aimed at increasing municipal infrastructure investments, leveraging local resources and involving a wide range of social partners. The study finds that the policy results in significantly increased numbers of both workers and plants, showing a positive correlation between funding intensity and the increase in plants.

However, many other scientific studies demonstrate that place-based policies do not achieve their primary objectives. For example, Neumark and Kolko (2010) which, using data at the plant level and geographical mapping techniques, shows that subsidies to Enterprise Zones in California have not led to a change in the occupational structure of low-wage workers, thus failing the primary objective of changing the composition of employment in those areas. In a similar work, the same authors, based on case study data collected from EZ administrators, reach the same conclusions although with more nuanced results that reflect variability in the implementation and local effects of EZ (Kolko and Neumark, 2010). Gibbons et al. (2021) show that a programme aimed at improving the quality of life in disadvantaged areas of the UK[6], despite large public investments, has had minimal results in terms of employment rates, as well as insignificant effects on job creation, generating no significant benefits for residents. Chen et al. (2019) assess the impact of China's "Development Zones" on Total Factor Productivity, using their mass closure as a natural experiment. The DiD results indicate that the end of these policies resulted in a consistent reduction in Chinese firms' TFP, with a geographically heterogeneous intensity, more pronounced in proximity to major agglomerated seaports.

In the context of European policies, Ciani and De Blasio (2015) show, that for the

[5]Objective 1 is a European Union programme that aims to promote economic development and reduce regional disparities in the most disadvantaged areas. This programme, linked to the Structural Funds, is aimed at regions with a GDP per capita lower than 75% of the EU average, making them eligible to receive funds directed at supporting economic growth, infrastructure development and job creation.

[6]The Single Regeneration Budget.

southern Italian regions, the EU Structural Funds relating to the 2007-2013 cycle, had a very limited impact on employment, as well as a null effect on economic growth, on the reduction of depopulation and the growth of real estate values. Albanese et al. (2023) report how the EU funds allocated to the Italian regions under Objective 1 may not have achieved their primary task of reducing inequalities. Indeed, the authors demonstrate that income inequalities in the treated areas decreased with the end of the funding programme. This reduction is determined by the shift of higher incomes towards the centre of the distribution, indicating that financial support may have favoured higher-income earners over those most in need. The study also highlights that a low quality of institutions may have accentuated inequalities through rent-seeking mechanisms activated during the funding period.

At the firms level, Bronzini and De Blasio (2006) observed that state public subsidies, aimed at companies in economically depressed areas in Italy, mainly led to an inter-temporal substitution effect of business investments and not to long-term benefits. This behaviour implies that subsidies do not generate new economic activity, but only a temporary alteration in the timing of investments already planned. Andini and De Blasio (2016), instead, examine the effectiveness of "Program Contracts in Italy", a place-based policy that financed investment projects for private companies in areas with low economic performance. They show a very negative impact of the programme, with benefits limited only in the single municipal areas treated. This effect is, however, mainly attributable to spatial relocation, or negative spillovers, in which economic entities move their plants from nearby untreated areas to treated ones in order to access the benefits offered by the policy.

Accetturo and De Blasio (2012), in a study on bottom-up place-based policies (which are less explored in the literature, compared to interventions directed at the industrial sector and the labour market), analyse the effects of the Italian "Territorial Pacts" on the number of municipal plants and employment. This policy is distinguished by its bottom-up participatory local development model, involving neighbouring municipalities, eligible for EU cohesion funds, and civil society actors, in the creation of development plans for disadvantaged areas. Despite the innovative approach of this policy, the study does not highlight significant improvements in employment growth or the number of plants between the treated municipalities and the control group, indicating the ineffectiveness of the policy in generating lasting socioeconomic benefits.

Despite the univocal evidence, both positive and negative, shown so far, the major part of the literature on place-based policies shows discordant results. In a very relevant econometric work, Kline and Moretti (2014a) analyse the spatial effects of one of the most important place-based programmes aimed at regional development approved in the United States: the "Tennessee Valley Authority" (TVA). The authors highlight that the policy had generated positive effects on the regional economy in agricultural employment, but only in the short term, and positive and significant impacts on national manufacturing employment in the long term, through agglomeration economies. However, local benefits have not translated into national economic improvements. Consequently, the positive spillover effects, localised within the target region, are offset by significant losses in other regions. Again for the USA, Greenstone et al. (2010) quantify the agglomeration spillovers between production sites, relative to the opening of large production plants. The authors verify, through microeconometric estimates, that the opening of a large plant increases the TFP of existing firms in the nearby counties, compared to that of the eligible counties. However, the opening of a new large plant in the treated counties determines an increase in labour costs that prevents profits from growing at the same rate as productivity increases. Busso et al. (2013) observe that the participation of specific target areas of the United States in EZ programmes has led to significant improvements in the labour market, with increased employment and wage growth, moderate increases in property values and rents, and minimal improvements in local demographics. Adopting a comparative perspective, Zheng et al. (2017) document that the growth in plant employment and wages, resulting from China's investment programmes in Industrial Parks, had generated positive economic and demographic spillovers, including the construction of new homes and businesses near the areas treated. However, these benefits are not spatially uniform. Specific factors, such as human capital, foreign direct investment, and industrial agglomeration, led to significantly greater gains.

In a spatial econometric study Ehrlich and Seidel (2018) examine the long-term effects of place-based subsidies in the West German Zonenrandgebiet[7]. The results show persistent positive effects on economic density, even post-subsidies. However, doubts arise about the overall effectiveness of the policy, as a significant part of economic activity results from negative spillovers and spatial reallocations rather than from net new growth.

In a significant body of causal econometric work Becker et al. (2010, 2012, 2018) based on EU structural funds earmarked for convergence of Objective 1 regions, shows mixed results. In particular, the authors, analysing the first three EU programming cycles (from 1989 to 2006), find a positive effect of the funds on the growth of per capita income in the regions treated, but no significant impact on employment (Becker et al., 2010). In another work, they select only the funding cycles from 1994 to 2006 and obtain similar results on per capita GDP. However, they note that, in many areas, the treated level of transfers is higher than the level that maximises aggregate efficiency, suggesting suboptimal management that compromises the EU's convergence objectives (Becker et al., 2012). In a study extended to a broader period of EU programming (1989-2013) that includes the economic crisis that began in 2007, the authors show positive effects on GDP per capita only during the treatment period. They also verify that in periods of crisis, the effects of the funds are weaker (Becker et al., 2018).

Still in the EU context, Cerqua and Pellegrini (2018) analyse the impacts of ESF on the economic growth of European regions, exploiting the intensity of the treatments

[7]The Zonenrandgebiet is an area, close to the Iron Curtain, that received subsidies to mitigate the economic challenges of isolation from Eastern markets during the Cold War.

received. Although the average results are positive, the authors observe diminishing returns. The concave relationship between financing intensity and economic growth suggests the existence of an optimal level of financing beyond which further increases do not contribute significantly to growth.

In the Italian context, Albanese et al. (2021) examine the effectiveness of the EU Structural Funds for the promotion and growth of TPF in the Italian Mezzogiorno, between 2007 and 2015, finding positive impacts of the funds. However, such impacts seem to be greater in areas characterised by better institutional quality and a higher population density. Cerqua and Pellegrini (2014, 2017) analyse the effectiveness of the main Italian programme of subsidies to firms in disadvantaged areas: Law 488/92[8]. They find that subsidies had positive and significant effects on firms' capital, turnover and employment in Southern Italy, but a negligible impact on productivity. Although transfers stimulated new investments, these did not lead to further self-financed investments (Cerqua and Pellegrini, 2014). Furthermore, subsidies have had a negative impact on TFP, with negative spillovers penalising non-treated firms, increasing competition in the labour market (Cerqua and Pellegrini, 2017). In a more recent empirical work, Ciani et al. (2024) analyse the effects of subsidies on the long-term hiring of unemployed people in Southern Italy. The findings reveal a significant increase in employment among the long-term unemployed, as well as a tax benefit associated with increased employment, that exceeded the costs of the treatment. However, following the end of the subsidy, a sharp decline in the probability of employment was observed.

As discussed, the increasing importance of place-based cohesion policies in the institutional context, together with the variable effects observed, highlights the importance for social scientists to assess their effectiveness. Despite the lively debate on these policies, the current literature, to the best of our knowledge, does not include counterfactual works examining the causal effects of the National Strategy for Internal Areas (SNAI) - the main bottom-up policy implemented on the territory in Italy - on crucial socio-economic variables, such as depopulation, employment and the number of production plants. Indeed, to date, the available studies have focused exclusively on the use of the SNAI municipal classification (illustrated in the next Section, 1.3), or on the theoretical aspects that established it, to try to explain socio-economic phenomena and trends, such as demographic decline (De Renzis et al., 2022; Sonzogno et al., 2022), territorial marginality (Vendemmia et al., 2021), income inequality (Gallo and Pagliacci, 2019; Guzzardi and Morelli, 2024), and local entrepreneurship (Mastronardi et al., 2020). In this sense, our work aims to fill this gap, offering the first econometric contribution that explores these dynamics, significantly enriching the corpus of existing studies.

[8]Law no. 488 of 1992 was the main policy instrument to reduce territorial disparities in Italy during the decade 1996-2007.

1.3 Institutional Setting

Since the Treaty of Lisbon, part of the EU policies is specifically aimed at hindering multidimensional poverty and - between and within - inequalities in the different territories and disadvantaged countries (Atkinson and Piketty, 2007; Alvaredo et al., 2013; Liberati and Resce, 2022). In this regard, the role played by the European Rural Development Policies is gaining momentum (Shucksmith et al., 2005; Crescenzi and De Filippis, 2016). For the purpose of targeting within the larger EU rural development framework, in 2012, the Italian Government introduced the definition of "Inner Areas" - territorial areas burdened by social, economic and environmental issues that have persisted for many decades (Barca et al., 2014). In particular, Inner Areas are characterised by a distressing economic and productive weakness in terms of employment and entrepreneurship, low levels of income and wealth and relevant disparities in infrastructural and essential services, concerning the centres. Furthermore, high depopulation rates, youth emigration, limited birth rate and ageing inhabitants are observed (Barca et al., 2014; Gallo and Pagliacci, 2019; Vendemmia et al., 2021; De Renzis et al., 2022; Sonzogno et al., 2022). Despite their economic disadvantages, according to Barca et al. (2014), Inner Areas have capabilities that can be exploited as they are generally located in rich environmental and cultural systems that have resources.

For this reason, in 2014, the Internal Areas were organised in a specific and general cohesion policy framework focused on local development and citizenship rights, i.e. the SNAI. This policy is accepted in the European programming cycle 2014-2020 (Barca et al., 2014; European Commission, 2014). The main objective of the SNAI is to reduce negative demographic trends, depopulation, and emigration while promoting local wellbeing and resilience. In this sense, the SNAI represents the most important and organic policy approved in Italy in the fight against depopulation, since other regulatory schemes have only concerned specific areas and territorial and regional contexts, sectors, and fields of intervention. These include, for example, the National Strategy of Green Communities of 2015, the Law on Small Municipalities of 2017, and specific regional laws [9].

SNAI is conceived as a place-based bottom-up strategy, favouring institutional cooperation at different levels and promoting participation and collaboration with all the economic actors, both public and private (Barca et al., 2014). In order to define the potential beneficiaries among the approximately 8,000 Italian municipalities, the policy first defines as non-beneficiaries the "Centres", that is, municipalities, or a set of neighbouring municipalities, endowed with: (i) schools offerings education from primary to secondary level; (ii) a hospital facility for emergencies with rescue services and general medicine functions; (iii) at least one regional railway station. Consequently, the Inner Areas, i.e. the potential beneficiaries, are classified on a residual basis, based on the accessibility to these services, or rather on the distance expressed in minutes travelled by car, from the

[9] For a review of these policies see (Cipolloni, 2021).

remaining municipalities to the nearest centre (Barca et al., 2014). In order to be classified as inner, a municipality must be at least twenty minutes away from a centre. Specifically, based on travel time, SNAI divides the Inner Areas into three categories (see Table 1.1): 1) intermediate areas, for municipalities 20 to 40 minutes away from the nearest service hub (D); 2) peripheral areas, for those which are 40 to 75 minutes away (E); ultra-peripheral areas, for municipalities located more than 75 minutes away (F).

Table 1.1: SNAI Classification

Centres
A - Hub pole B - Intermunicipal poles C - Urban belts
Inner areas
D - Intermediate area distance from the poles between 20 and 40 minutes E - Peripheral areas distance from the poles between 40 and 75 minutes F - Outermost areas distance from the poles between 40 and 75 minutes

Source: Author's processing on Department for Cohesion Policies.

The Inner Areas cover a large proportion of the Italian territory, accounting for most Italian municipalities, i.e. 53%, with a residential population of 13.5 million inhabitants, accounting for more than 20% of the total. Figure 1.2 presents a view of the distribution of Inner Areas throughout the Italian territory (Barca et al., 2014).

In 2015, to optimise the use of SNAI financial resources, the "Technical Committee for Inner Areas"[10] identified the so-called project area (or pilot areas), on which the local development practices underlying the policy would be implemented. The process of defining and selecting these areas developed in different and complex phases (Dipartimento per le politiche di coesione, 2016, 2017, 2018). Given the aim of the policy to promote institutional cooperation, eligible municipalities were encouraged to join consortia to advance their applications. This was a fundamental step as, for historical reasons dating back to the end of the Middle Ages, the municipalities in Italy, especially the smaller ones, were characterised by parochialism, that is, by fights and envy, between neighbouring municipalities. The subsequent phases involved a quantitative and qualitative assessment, and a diagnostic step of a socioeconomic and demographic nature. In particular, attention was focused on specific structural characteristics connected to delays in local development (see Table A4, which reported the indicators considered), on the data relating to the

[10]The Technical Committee for Inner Areas is coordinated by the Department for Territorial Development and other public and territorial actor.

requests of the municipal groups and on a subsequent field verification in these areas by the Committee. This investigation process led, in 2017, to the identification of seventy-two pilot areas (see Figure 1.1), which included 1,077 municipalities (more than 25 percent of the municipalities identified as Inner Areas) and approximately 2 million citizens. These municipalities, bordering each other, are mostly municipalities classified as peripheral or ultraperipheral, located in the southern regions, in mountainous sites, which suffer from ageing and a significant loss of resident population, as well as low-income levels.

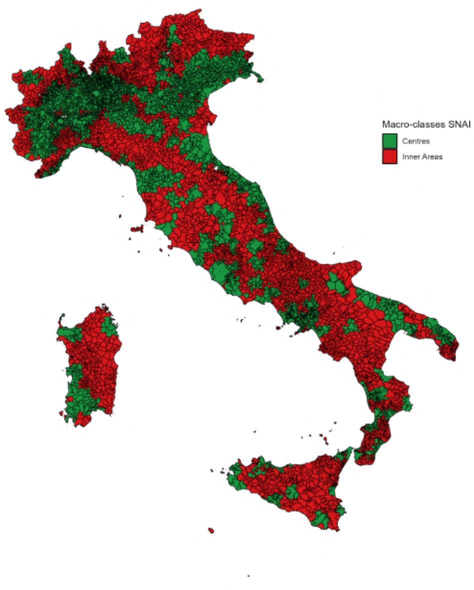
Once the individual pilot areas have been approved, the municipalities within them are called upon to take action and collaborate by following a detailed and complex process for the approval of "Framework Programme Agreements" (PFA), technical planning documents that summarise all the projects to be implemented, including financial details (Dipartimento per le politiche di coesione, 2017). So, the closure of the PFA makes the individual project areas eligible for treatments.

In this regard, it is clear that the treated units over the 2018-2020 period are subject to non-random selection and this is not comparable with non-treated units, even if eligible for the policy. However, two major points are worth noting. First, by construction, treated units are necessarily similar and comparable. Second, and more importantly, the treatment timing can be considered random. In particular, there are time lags and delays in obtaining benefits, determined by the imperfect exchange of information, between local and central authorities, the limited scarcity of available resources, as well as the feasibility and monitoring of interventions. These problems led to an assignment of treatments that followed an asynchronous trend, concerning the PFA closure process, and was not spatially uniform at a territorial level. In this sense, the benefits were granted with priority being given to the areas that presented projects with the best timetables and with more structured inter-municipal cooperation (Dipartimento per le politiche di coesione, 2016, 2017, 2018, 2019, 2020, 2021). Indeed, in 2020, the last year of observation in the sample, even though all the PFA had been elaborated, concluded, and approved, a non-negligible part of the projects had not yet been financed.

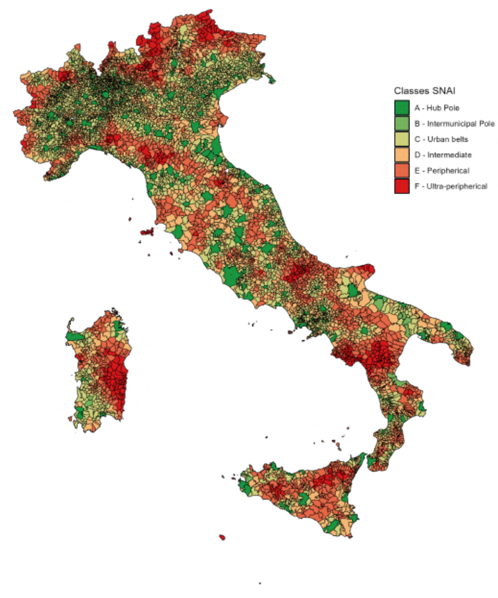
From a financial point of view, SNAI is made up of national resources, allocated with the Budget Laws, and community funds, such as ERDF, ESF, and FEASR, as well as other public and private resources. The overall allocation for the programming of interventions in the municipalities of the seventy-two pilot areas, at the end of 2020, exceeded 1 billion, 150 million euros.

Figure 1.1: Spatial Distribution of Municipalities Between Centers and Inner Areas According to SNAI

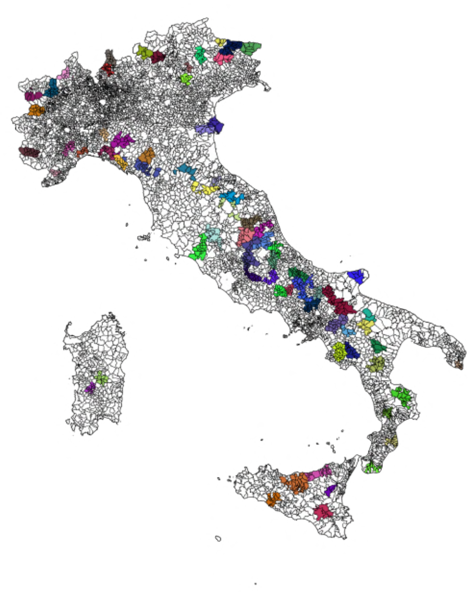
(a) Inner Areas and Centers



(b) SNAI Classes



(c) Pilot Areas



Source: Author's processing on Department for Cohesion Policies data.

Table 1.2 and Figure 1.2 report the information relating only to the municipalities of the financed pilot areas (Dipartimento per le politiche di coesione, 2019, 2020, 2021)[11]. In particular, by December 31, 2020, only 19 pilot areas, corresponding to 269 municipalities, had received payments, for an amount exceeding 29 million euros, over the two-years 2018-2020. It is worth noting that the amount of payments made is remarkably modest. The main limitation of the overall SNAI policy is that, at the end of 2020 (the last data used for the empirical analysis), the amount of money spent for the policy enactment was less than 10 percent of the total funds' programmed amount (29 million spent, 390 million programmed in these areas). On average, the treatment received by each municipality ranged from around $70k$ € in 2018 to $110k$ at the end of 2020. Despite the small level of investment that highlights an overall limitation of the SNAI policy framework, as the treated municipalities were few and small, the payments can make a difference in the local economy (the average resident population of the municipalities in these areas is just over $2,500$ residents). In the pilot areas, the payments are about 950 € per inhabitant, which is higher than what a municipality spends on average for relevant services, such as local police (Bucci et al., 2023). Furthermore, the figures increase in the future, as the total funds' programmed amount to about 390m €.[12].

[11]For a more complete detail of the treatments in the pilot areas see Table A5.

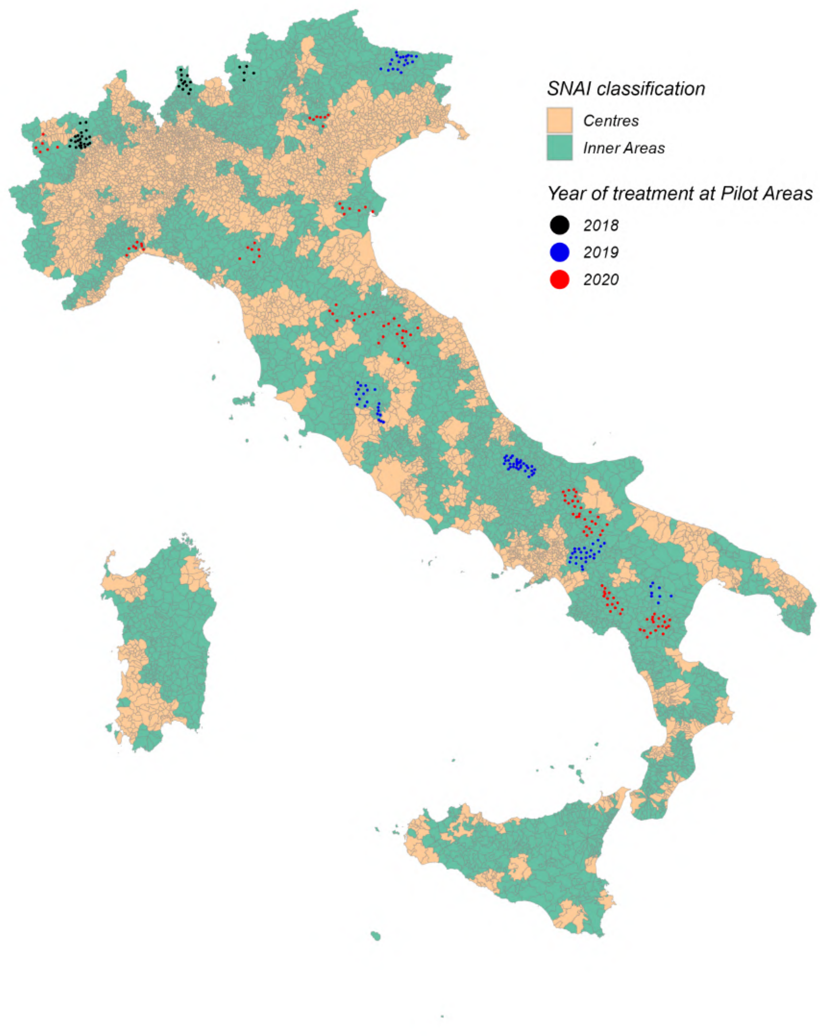
[12]Our study does not consider the years after 2020 to avoid two main confounding factors, which made the data on financing and controls of the "Annual reports on the implementation of the policy" unavailable for the years after 2020, namely: the institutional changes that involved the SNAI for the expiry of the European cohesion programming for the years 2014-2020, and the outbreak of COVID-19 and its policy consequences. Undoubtedly, the COVID-19 pandemic has worsened social, economic, and territorial problems (Carillo and Jappelli, 2022; Cerqueti et al., 2022). Lockdowns and market closures have brought the entire world economy to its knees. The Italian GDP, in 2020, recorded a loss of 8.9% (ISTAT, 2021). Poverty and inequalities, especially in health and education, have significantly increased (Brunori et al., 2021). The EU's response to the recession has been powerful. With the European funds of the Next Generation EU (NGEU), the Community institutions had intended to promote robust and sustainable economic growth throughout the continent (Cerniglia et al., 2021). The Italian National Recovery and Resilience Plan (PNRR) detailed multi-annual structural interventions, organised in a multilevel structure, which involved local and regional administrations (Maranzano et al., 2021). It paid considerable attention to the problems present in the Inner Areas. It refinanced the Strategy with significant resources, which were combined with national funding approved during the most acute phase of the pandemic emergency (Dipartimento per le politiche di coesione, 2021). In addition to the funds for the recovery for the strengthening of the SNAI, in the coming years, other resources will be assigned, as part of the new European programming of cohesion policies, 2021-2027. The EU, therefore, recognises the importance of SNAI as an economic and cohesion policy capable of reversing negative trends. In this sense, COVID-19, for SNAI, represented a sort of watershed between the new and old European programming and which was nationally oriented for marginal areas (ISTAT, 2022).

Table 1.2: SNAI Financial Progress: Treatments in the Pilot Areas (2018–2020)

Treatment Periods	**Total Scheduled**	**Payments Times**		
		2018	**2019**	**2020**
Total Scheduled and Treatment	389,858,749	2,805,298	7,968,357	29,243,973
Per municipality	1,449,289.03	70,132.46	54,954.19	108,713.66
Per inhabitant	948.39	45.34	55.30	72.48
Number of Pilot areas treatment	**19**	**3**	**8**	**19**
Number of inhabitants in Pilot Areas	**269**	**40**	**145**	**269**
Total number of inhabitants in Pilot Areas	**-**	**65984**	**233395**	**565428**
Average inhabitants in the Pilot Areas	**-**	**2251**	**1868**	**2520**

Source: Author's processing on Department for Cohesion Policies data.

Figure 1.2: Spatial Distribution of Pilot Areas Treated According to SNAI (2018-2020)



Source: Author's processing on Department for Cohesion Policies data.

The strategy divides the recurring interventions in the pilot municipalities into two major pillars: essential services, which focus on mitigating the distance from central services, and local development, focused on overall economic development.

Many welfare practices and initiatives have already been carried out during the first years of the implementation of the SNAI. Interventions have been carried out to improve citizens' quality of life, enhancing the potential that the territories offer.

The first resources, committed to strategic projects in the pilot areas, concern the areas of intervention attributable to the economic and productive sectors that favour development and the local economy, such as agriculture, land care, the efficiency of public administration, infrastructure, tourism, waste materials and separate disposal collection, etc (Dipartimento per le politiche di coesione, 2019, 2020, 2021). Some municipalities in the pilot areas have used SNAI funds to improve energy efficiency in public lighting networks and municipal offices[13].

Moreover, the interventions implemented by local administrations and actors have contributed to increasing and enriching the offer of essential services for citizens, in particular in the socio-health, education and training, and mobility and transport sectors (Dipartimento per le politiche di coesione, 2019, 2020, 2021). Practically, in the sectors of healthcare and assistance to people with SNAI resources, facilities have been opened for the elderly and sick people who require long-term care. For example, developing specific infrastructures and activities have created new entrepreneurial opportunities, such as the establishment of specialised healthcare and assistance units for specific pathologies.[14]. Furthermore, innovative professionals have been assigned to carry out territorial proximity work that is important for the survival of local communities, such as community nurses and midwives. The education sector is also at the centre of other initiatives. In many areas covered, projects have been launched for the modernisation and redevelopment of existing buildings with advanced digital technologies. These offer students, in peripheral municipalities, greater learning opportunities, reducing school dropout and emigration levels to that of more populated cities. Similarly, in the transport sector, projects aimed at sustainable mobility have been launched and tested, such as on-demand transport systems for workers and students, in addition to the construction of cycle paths.

So, although SNAI funds are not specifically intended for the establishment of new businesses, given the nature of the funding and the variety of feasible projects, the local interventions can increase economic activity and incentivise entrepreneurship in the target areas, as they involve policy actions that necessitate the presence of companies specialising in various sectors.

[13]For further information, read this document: https://www.agenziacoesione.gov.it/wp-content/uploads/2020/12/Basilicata-Montagna-Materana.pdf.

[14]Further information on some specific projects are publicly available: https://www.agenziacoesione.gov.it/wp-content/uploads/2020/12/Basilicata-Montagna-Materana.pdf.

1.4 The Dataset

Table 1.3 reports descriptive statistics broken down by Centres and Inner Areas. In turn, the latter are divided in relation to the SNAI treatment, i.e. into municipalities eligible for treatment and municipalities treated in the three cohort years analysed (2018-2019-2020). The variables are gathered into four groups obtained from three different sources and merged using the unique municipality identifier provided by the Italian National Institute of Statistics (Istat). The first group is intended to capture economic conditions (source: Italian Ministry of Economics and Finance); the second group captures demographic conditions (source: Istat); the third captures institutional characteristics (source: registry of local and regional administrators of the Ministry of the Interior); and finally, the fourth group pertains to the territorial morphology of the Italian municipalities (source: Istat). It can be noted that the inner municipalities show poor (average) performance on many of the indicators studied. Especially in the treated municipalities, the obtained results reveal the greater pervasiveness of multidimensional deprivations which negatively affect the quality of life and the wellbeing of the territories, since the latter areas are characterised by high socioeconomic criticalities, as well as by differences in the composition and in the structure of institutions and local administrations. Notably, these differences are made evident by the average number of *plants* (700 units on average more in the Centres than in the Inner Areas), and by the level of income. Furthermore, despite encompassing a greater *surface* area, the Inner Areas have, on average, a *population* which is a quarter of that of the Centres, negative net immigration flows, and a higher share of population *over 65* years, highlighting the ageing of municipalities as well as a higher proportion of the population that is out of the labour market. These figures, capture the main problems afflicting the Inner Areas, namely depopulation and abandonment. This phenomenon is also partly attributable to the fact that these municipalities are mainly located in mountain areas, far from industrial and commercial centres and from more accessible logistics and distribution centres, as evidenced by the *altitude*. Other noteworthy characteristics that emerge from the table concern the lower proportion of *female mayors* and the greater proportion of municipalities administered by *civic lists* for the municipalities of the Inner Areas. These data show a weak capacity to innovate and adapt its organisational and governance structure. As regards the remaining institutional characteristics, such as *age of mayor*, *intensity of graduated mayors* and *average age of municipal councillors*, the two groups are similar. In this sense, it is, in fact, noticeable in the last three columns of Table 1.3, that there are fewer evident gaps between eligible Inner Areas and treated municipalities.

Overall, a wide heterogeneity in economic, demographic, and institutional features can be observed, calling for support and public policies aimed at generating development in the most marginal municipalities and reducing inequalities between urban and rural areas.

Table 1.3: Descriptive Statistics for Centres and Inner Areas (Eligible and Treated in 2018, 2019, and 2020)

	Centres	**Inner Areas**			
		Eligible	Treated in 2018	Treated in 2019	Treated in 2020
Plants	934.25	218.37	134.67	120.52	198.11
	(5704.45)	(416.58)	(149.17)	(206.75)	(215.83)
Average Income	23325.53	20264.14	22066.05	18974.59	19231.03
	(3161.25)	(2854.08)	(2034.38)	(1993.40)	(2147.06)
Demographic					
Population	11724.21	3264.73	1636.38	1746.88	2828.89
	(57077.89)	(5713.51)	(1521.16)	(2171.11)	(2760.52)
Net immigration flow	4.22	-4.49	0.23	-4.77	-7.07
	(148.98)	(30.80)	(13.09)	(15.15)	(18.83)
Over 65	21.69	25.34	22.10	28.80	26.92
	(4.27)	(5.68)	(3.10)	(6.57)	(4.64)
Institutional characteristics					
Mayor's age	52.05	52.18	48.79	52.94	51.72
	(10.64)	(10.89)	(12.40)	(10.70)	(10.57)
Mayor's education (degree)	0.45	0.43	0.24	0.39	0.48
	(0.50)	(0.49)	(0.42)	(0.48)	(0.49)
Female mayor	0.16	0.12	0.14	0.10	0.09
	(0.37)	(0.32)	(0.34)	(0.29)	(0.28)
Average age of council	47.36	46.38	45.77	46.24	45.31
	(4.54)	(4.91)	(4.39)	(4.52)	(4.41)
Civic list	0.60	0.60	0.33	0.61	0.58
	(0.49)	(0.49)	(0.47)	(0.48)	(0.49)
Geographical variables					
Altitude	208.51	476.31	752.72	655.14	602.68
	(181.56)	(305.49)	(434.28)	(226.49)	(280.91)
Surface	30.86	43.52	54.39	49.74	69.14
	(51.98)	(46.27)	(50.11)	(47.14)	(50.56)

Note: average values at municipal level and standard deviation in parenthesis. *Plants* = number of plants; *Average income* = inhabitants' average income; *Population* = number of inhabitants; *Population density* = share of inhabitants; *Over 65*= percentage population over 65; *Mayor's age* = age of the mayor; *Mayor's education* = 1 if mayor has a degree; *Female mayor* = 1 if mayor is female; *Age of council* = average age of municipal council members; *Civic list* = 1 if the municipality is administrated by a civic list; *Altitude* = altitude in meters; *Surface* = surface in km^2.

1.5 Empirical Strategy

1.5.1 Recent Developments in the Difference-in-Difference

Differences-in-difference (DiD) is an estimator which is widely used by applied researchers (Cerulli et al., 2015; Cerulli and Ventura, 2019). Under the assumption of parallel trends between treated and non-treated units and no anticipation effects, this estimator can consistently estimate the causal effects in non-experimental settings. However, a recent body of literature has shown that the estimator suffers from serious drawbacks under staggered entry. When units enter into treatment at different points in time, i.e. by cohorts, the canonical two-way fixed effects (TWFE) coefficients may even take the opposite sign with respect to the true one, essentially due to negative (or at least incorrect) weightings attached to each cohort, according to Callaway and Sant'Anna (2021) and Sun and Abraham (2021), the most prominent among the authors. Moreover, the inclusion of treatment lags and leads in the TWFE regression, as a test for pre-trends, can be very misleading due to heterogeneous treatment effects across cohorts. Indeed, the coefficients of lags and leads can be contaminated by effects from other periods and, consequently, the test is inconclusive. Among the different solutions offered by the new strand in the literature, Sun and Abraham (2021) propose a method that corrects the incorrect weightings attached by the TWFE to each coefficient in such a manner as to retrieve unbiased estimates of the average treatment effect for the treated (ATT) at different relative times, i.e., at different treatment times for each cohort. The distribution of these coefficients is commonly referred to as event study analysis and is satisfactorily reported in a plot that gives an immediate impression of the policy effects unfolding over time.

The policy we are considering falls exactly within this scenario, as its implementation hitherto has occurred in three subsequent cohorts, 2018, 2019, and 2020. Therefore, taking advantage of the recent advances in econometric theory, we estimate the effects of the policy by applying the estimator proposed by Sun and Abraham (2021). Essentially, this method is based on the idea of using the proportion of cohorts as weightings, because these are more interpretable than the weightings implied by the TWFE and sum up to unity. The resulting estimates are robust to treatment effect heterogeneity.

The identifying assumptions are the same as the well-known assumptions of DID, such as (i) parallel trends in baseline outcomes; (ii) no anticipatory effects prior to treatment. The main building block of the estimates is the cohort average treatment effect on the treated (CATT), defined as the cohort-specific average difference in outcomes relative to a suitable comparison group. It follows that the entire procedure can be synthesised in three steps. First, estimate each CATT using an interacted TWFE regression, where interactions take place between cohort dummies and relative time dummies. Estimate the weightings and finally, average the CATT's using the weightings. A more formal description of the estimator is given in the next section and for a complete guide we refer the interested reader to Sun and Abraham (2021).

As regards the comparison group, one can either use the never treated units or the last cohort to be treated. In order to make the treated and the non-treated group as similar as possible, we first restrict our attention to eligible units only, and then we choose the last treated cohort as the most similar. In fact, as noted by Sun and Abraham (2021), never-treated are likely to behave differently from last-treated; for this reason, and in view of the fact that adherence to pilot areas depends on an initial proposal from municipalities that results in selection bias, we choose the last treated cohort, i.e., 2020, as the control group[15]. In our specific setting, the outcome variable is the number of plants in each municipality and the parallel trend is plausible for at least two reasons; (i) treated and last treated are chosen on the basis of common characteristics; (ii) the municipalities in these cohorts are affected by serious and structural development problems, due to being located in mountain regions and far from vital connections to market outlets. These time-invariant characteristics strongly affect the outcome variable and cannot be changed, whence the need for intervention. In other words, in the absence of intervention, all the cohorts would keep following (unfortunate) analogous paths in entrepreneurship. As far as anticipatory effects are concerned, we claim that they are very unlikely to hold again on the basis of two structural features. Firstly, due to historical and cultural heritage, Italian local realities are characterised by strong inertia, even inactivity, and it is difficult to think of a (positive) reaction to the policy before its actual implementation (Barca et al., 2014; Mastronardi et al., 2020; Vendemmia et al., 2021; De Renzis et al., 2022; Sonzogno et al., 2022). Simply, in order to provide some figures supporting this claim, consider that only around 40% of the mayors of the treated municipalities hold a degree and only 10% of mayors are females, pointing to a conservative cultural heritage. Secondly, the policy was completely new and to a certain extent, its effects were unexpected. In environments characterised by resistant attitudes to change and substantial uncertainty about the benefits, it is unlikely to expect agents to behave strategically, such as by anticipating the unknown benefits.

Nonetheless, uncertainty about future benefits raises strong suspicions of heterogeneous causal effects across cohorts where the impact of the policy unfolds over time as the agents, and therefore the cohorts, become acquainted with the policy and its potential benefits. In other words, there are reasons to argue that treatment heterogeneity would be in place, invalidating the TWFE estimates. Indeed, cohorts may differ in their covariates, which affects how they respond to treatment. For instance, if treatment effects differ with municipal average income or the share of poverty at the municipal level, we will have heterogeneous effects. Furthermore, treatment effects may vary across cohorts due to cyclical macroeconomic conditions. Nonetheless, these sources of heterogeneity are still compatible with our parallel trend assumption, which only rules out selection into treatment timing based on the evolution of the baseline outcome.

[15] When the last treated is used as control group one needs to drop periods beyond the date at which the last cohort was first treated, i.e. 2020 in our case.

1.5.2 The Model

This section describes the Sun and Abraham (2021) model, focusing only on those parts of the model that strictly concern our configuration. Let us consider a panel data set up with $T+1$ time periods and N units. Let us assume to observe the outcome variable Y_{it} and the binary treatment status D_{it} with $i = 1, \ldots, N$ and $t = 1, \ldots, T+1$. The treatment status takes the value 1 if and when unit i is treated. We also assume the treatment status to be of the absorbing type, i.e. once a unit is treated, it remains treated in the following periods. It follows that it is possible to group units by cohorts, namely according to the first treatment time, $E_i = min\{t : D_{it} = 1\}$, and denoting never treated units by $E_i = \infty$. Thus, units in the same cohorts e for $e \in \{0, \ldots, T, \infty\}$ are first treated at the same time $\{i : E_i = e\}$. As for the potential outcome, we define Y_{it}^e as the potential outcome in period t when unit i is first treated in time period e. By analogy, Y_{it}^∞ is the potential outcome under no treatment. In order to make identification possible the following assumptions are made:

Assumption 1: Parallel Trend in Potential Outcomes:

$$E\left[Y_{i,t}^\infty - Y_{i,s}^\infty \mid E_i = e\right] \text{ is the same for all admissible } e \text{ and for all } s \neq t \tag{1.1}$$

where:

- $Y_{i,t}^\infty$ is the potential outcome for unit i at time t if never treated;
- E_i is the treatment group identifier for unit i, where e represents the time the unit first received treatment. Units never treated are denoted by $E_i = \infty$;
- s is the reference time period for comparison;
- t is the target time period being analysed.

Assumption 2: Absence of Anticipation Effects:

$$E\left[Y_{i,e+l}^e - Y_{i,e+l}^\infty \mid E_i = e\right] = 0 \quad \text{for all admissible } e \text{ and } l < 0 \tag{1.2}$$

where:

- $Y_{i,e+l}^e$ is the potential outcome for unit i at time $e+l$ under the assumption the unit was treated at time e;
- $Y_{i,e+l}^\infty$ is the potential outcome for unit i at time $e+l$ if never treated;
- l is a time period before treatment ($l < 0$), ensuring there are no effects from anticipating the treatment.

Using the notation at hand and the event study setup, we can formally define the building block of even the study analysis, as the average of unit-level treatment effects at relative period l for units belonging to cohort $E_i = e$ as:

$$CATT_{e,l} = E\left[Y_{i,e+l} - Y_{i,e+l}^\infty | E_i = e\right] \tag{1.3}$$

where:

- $CATT_{e,l}$ measures the average treatment effect on the treated at time e observed l periods after treatment;
- e and l define specific cohort and time relative to treatment initiation.

In general, $CATT_{e,l}$ represents the ATT of units first treated at time e, l periods since treatment. Usually, l is referred to as the relative period $l \in [-e; T-e]$. Under the two assumptions above, Sun and Abraham (2021) propose to estimate a weighted average of $CATT_{e,l}$ with convex weightings given by the share of cohorts that have been treated at least for l relative periods. In addition, such a weighted average is normalised by the number of relative periods included in the estimate. Formally the estimator is given by

$$v_g = \frac{1}{|g|} \sum_{l \in g} \sum_{e} CATT_{e,l} Pr\left\{E_i = e | E_i \in [-l, T-l]\right\} \tag{1.4}$$

where:
- $|g|$ is the number of relative periods included in the estimate;
- $CATT_{e,l}$ is the Conditional Average Treatment Effect;
- $Pr\left\{E_i = e \mid E_i \in [-l, T-l]\right\}$ is the probability of being treated at time e given being in the cohort during the relative periods.

From a practical point of view, the estimator provided by Sun and Abraham (2021) can be implemented in three steps which aim at estimating the sample analogue of each component of (4) and to feed them back into the population equation. Hence, one must first estimate $CATT_{e,l}$ using a TWFE augmented by the interactions between relative period indicators and cohort indicators excluding at least one cohort indicator from the set C:

$$Y_{it} = \alpha_i + \lambda_t + \sum_{l \notin C} \sum_{l \neq -1} \delta_{e,l} \left(\nVdash \{E_i = e\} D_{it}^{l} \right) + \epsilon_{it} \tag{1.5}$$

where:
- α_i and λ_t are unit and time fixed effects, respectively;
- $\delta_{e,l}$ captures the interaction effects of treatment status across different cohorts and time periods;
- D_{it} is the binary treatment status for unit i at time t, which takes the value 1 if the unit has been treated at time t or any time before t and remains 1 thereafter, reflecting the absorbing nature of the treatment;
- ϵ_{it} is the idiosyncratic error term.

In 5 the excluded period is -1 and the estimated CATT's can be interpreted as the effects of the policy with respect to the previous relative time, e.g. $CATT_{e,l}$ is the average effect at relative time l with respect to $l-1$, for cohort e. The cumulative effect of the policy at time m can be obtained as $\sum_{e,l=0}^{m} CATT_{e,l}$. As a comparison group, if there are never treated we may set $C = \infty$, otherwise, we may set $C = \{max\{E_i\}\}$ i.e. the latest treated cohort and restrict the estimation sample to $t = 0, \ldots, max\{E_i\} - 1$. As a second step, it is necessary to estimate the weightings as the proportions of each cohort in the relevant period l.

The final step consists of aggregating the CATT's using the weightings from the previous steps. Formally:

$$\hat{v_g} = \frac{1}{|g|} \sum_{l \in g} \sum_{e} \hat{\delta}_{e,l} \hat{Pr} \left\{ E_i = e | E_i \in [-l, T-l] \right\} \tag{1.6}$$

where:

- $\hat{v}_g$ is the estimated average treatment effect across specified relative periods;
- $|g|$ is the number of relative periods included in the estimate;
- l indexes the relative periods since treatment within the set g;
- e represents the cohort first treated at time e;
- $\hat{\delta}_{e,l}$ is the estimated treatment effect for cohort e at relative period l;
- $\hat{Pr} \left\{ E_i = e \mid E_i \in [-l, T-l] \right\}$ is the estimated probability that unit i, belonging to cohort e, is observed in the relative period l, given that it is part of the data sample between periods $-l$ and $T-l$.

This estimator is referred to as an "interaction-weighted" estimator (IW) and is a consistent estimator of a weighted average of CATT's. The absorbing treatment state is actually the case of our empirical application because the treated municipalities are treated up to the end of the sample; however, the effect of ever being treated may still be interesting in many contexts, such as the effect of minimum wage on employment (Callaway and Sant'Anna, 2021) or the effect of hospitalisation on future labour earnings (Sun and Abraham, 2021). Moving from the theory to the policy under evaluation, our outcome variables are the net immigration flows per 1,000 inhabitants, and the number of plants per municipality. Our binary treatment status, D_{it}, takes the value of 1 starting from year t in which municipality i begins receiving SNAI funding. As discussed in Section 1.3, SNAI focuses on two main areas of intervention: essential services and local development. However, we do not distinguish between these two areas in our econometric analysis due to the lack of detailed municipal-level data. Additionally, we anticipate that improvements in services are less likely to occur in the short run. Nonetheless, we expect that, in the short term, both areas of intervention can foster entrepreneurship, which is a necessary condition for population growth.

In addition, in order to corroborate the results, we test the baseline model controlling for the following sets of covariates:

1. a first set is intended to capture the economic conditions of inhabitants, *eco*, containing the average income of inhabitants. This covariate is included to control for municipal heterogeneity in the economic conditions that may contribute to the creation of both the demand and the supply side (Gallo and Pagliacci, 2019). Data for *eco* are taken from the fiscal declaration data set available at the municipal level from the Italian Ministry of Economics and Finance;[16]
2. demographic, *demo*, consisting of the first lag of (log)population. We include

[16] https://www.finanze.gov.it/it/.

population to control for factors connected to the agglomeration, which is widely recognised as a key explanatory factor for the concentration of industrial activity by location theorists (Carlino, 1980). Data for *demo* are taken from Istat;[17]

3. institutional characteristics of the local government, *gov*, such as mayor's age, gender, and education (dummy for degree or more), the average age of municipal council members, and a dummy if the municipality is ruled by a civic list. We include *gov* since it has been largely shown that some institutional factors have an impact on the local creation of plants. Regarding the age of the mayor and councillors, Alesina et al. (2019) noted the tendency of younger politicians to behave strategically, increasing spending and obtaining more transfers from higher levels of government, and these factors can, in some way, affect our outcome variable. We also control for gender, as women in politics are usually more concerned about peoples' wellbeing, show higher cooperation and team working skills, and are less likely to engage in corruption, compared to their male counterparts (Chattopadhyay and Duflo, 2004; Hernández-Nicolás et al., 2018). Consequently, a higher female political participation may affect the policies implemented and the allocation of resources across different programmes (Funk and Gathmann, 2015). Education is used to control the local politicians' formal human capital which is likely to affect government quality (Geys, 2017). Data for *gov* are taken from the registry of local and regional administrators of the Ministry of the Interior;[18]
4. geographical variables, *geo*, altitude, and surface in square kilometers are included to account for the differentiation in logistic costs that can be generated by the heterogeneity in these features (Hesse and Rodrigue, 2004). Data for *geo* are taken from Istat.[19]

1.6 Results

This section presents the estimated effect of funding on long and short run outcomes, i.e. population structure and the number of plants, respectively. Section 1.6.1 reports the results for the net immigration flows per 1, 000 inhabitants showing no effect on these variables. Concerning the number of plans, the focus is on three aspects: the effects on the treated municipalities (Section 1.6.2), robustness checks (Section 1.6.3), and possible spillover effects on the neighbourhood (Section 1.6.4).

[17]Istat makes available the most recent official data on the population in Italian municipalities derived from surveys conducted at the Offices of Registry and Civil Status of Municipalities and from the Census of Population. Link: `https://demo.istat.it/`.

[18]This registry consists of the information on elected officials in municipalities, provinces, metropolitan cities and regions concerning biographical data, the list or group to which they belong or are connected, educational qualification and the profession exercised. Link: `https://dait.interno.gov.it/elezioni/anagrafe-amministratori`.

[19]Link: `https://www.istat.it/it/archivio/156224`.

1.6.1 The Effects on the Population Structure in the Treated Municipalities

As a first step, we assess whether the funding has generated any effect on the main outcome of interest, i.e., depopulation, using net immigration flows per $1,000$ inhabitants as a tool to investigate whether the policy's newly introduced economic opportunities act as a catalyst for generating fresh immigration currents, with the ultimate aim of mitigating the demographic decline (Basile et al., 2023).

Table 1.4 reports the estimates from the TWFE and IW estimator showing no statistically significant results. This evidence can be considered as quite an expected result considering the short time span of our analysis which does not allow us to capture the longer-run effects of the policy. The estimates have also been repeated by including covariates and the same result applies (columns 2-4). Accordingly, we move on to analyse possible short-run effects repeating the analysis on the number of plants.

Table 1.4: Event Study: Estimates of the Effect of SNAI on Net Immigration Flow

relative time	TWFE	IW (1)	IW (2)	IW (3)	IW (4)
-5	1.169	2.431	1.567	1.616	1.616
	(1.165)	(2.015)	(1.945)	(1.96)	(1.96)
-4	0.744	0.891	0.886	0.95	0.95
	(1.295)	(1.462)	(1.476)	(1.486)	(1.486)
-3	0.964	0.96	0.44	0.416	0.416
	(1.055)	(1.117)	(1.111)	(1.128)	(1.128)
-2	0.337	-0.981	-1.471	-1.365	-1.365
	(0.996)	(1.13)	(1.161)	(1.185)	(1.185)
0	0.354	-0.177	0.375	0.21	0.21
	(1.226)	(1.132)	(1.28)	(1.287)	(1.287)
1	0.443*	0.07*	-0.859	-0.953	-0.953
	(1.849)	(0.371)	(0.76)	(0.704)	(0.704)
N	2,087	1,788	1,746	1,714	1,714
eco			✓	✓	✓
gov				✓	✓
geo					✓

Note: Robust Std. Err. clustered at the municipal level in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. TWFE stands for Two Way Fixed Effects; IW - Columns (1)-(5) - stands for the interaction-weighted estimator. The bottom of the table indicates the sets of covariates included in each column. *eco*=average income of inhabitants;*gov* characteristics of the municipal council; *geo* geographic characteristics.

1.6.2 The Effects on Plants in the Treated Municipalities

Table 1.5 reports the results from the estimation of the TWFE and the IW estimator showing substantial differences between the two. Considering the TWFE pre-treatment coefficients, i.e., from -5 up to -2, if we were confident that the TWFE provided valid inferences, we would conclude that the anticipatory effects condition was violated. Fortunately, recent developments have proven that test to be invalid. The estimates from the IW estimator tell another story.

The IW estimator signals a positive effect of the policy since its inception. The coefficients associated with the treatment are positive and significant both in the year the policy was introduced and in the following year. Thus, there is a positive and persistent increase in plants in treated municipalities; the average treatment effect over the entire treatment period is equal to 4.23 with a Pval=0.001 and the cumulative effect is equal to 8.46 with an associated Pval=0.002. These results can be decomposed by looking at the CATT's. Interestingly, significant results already appear to be accruing to municipalities in the second calendar time period. Indeed, the fourth column of the table, $\hat{\delta}_{2018,l}$ reports the CATT for the 2018 cohort and no statistical effects at relative time 0 corresponding to calendar time 2018, the first year the policy was ever introduced. After one year, $l = 1$, the policy produced positive and significant effects for this cohort corresponding to the calendar year 2019. Symmetrically, the cohort first treated in 2019 exhibits positive and significant effects at $l = 0$, i.e. in 2019 and very close in magnitude to those of cohort 2018 at $l = 1$. This joint reading of the effects of the policy unfolding over time is made immediate by a visual representation shown in Figure 1.3.[20] The upper left picture reports the estimates from the IW column, along with 95% confidence interval, while the remaining two pictures report the ATT's in columns (4) and (5) of Table 1.5 respectively, in the upper right and lower left panel.

These results are in line with the previous literature providing evidence of a positive effect of the cohesion policies (Giua, 2017; Becker et al., 2018; Fattorini et al., 2020). Of course, it seems plausible that both the two main areas of intervention (essential services aimed at mitigating the distance, and local development aimed at stimulating the economy) have the potential to boost local entrepreneurship or, at the very least, to alleviate the decline in business establishments. Improving essential services can enhance the quality of life over time, making Inner Areas more attractive to potential entrepreneurs. Better connectivity also allows local businesses to access broader markets, suppliers, and networks. In the short term, however, the immediate impact on entrepreneurship is likely to be focused on businesses directly connected to essential services and their related activities. This could result in a rapid increase in demand for service-related businesses and greater appeal for small-scale service providers. On the other hand, efforts to promote local development have a stronger potential to provide an accelerated economic boost through local projects. The

[20]Recent literature has argued that event study plots in a non staggered setting cannot be compared to dynamic TWFE plots, see Roth (2024).

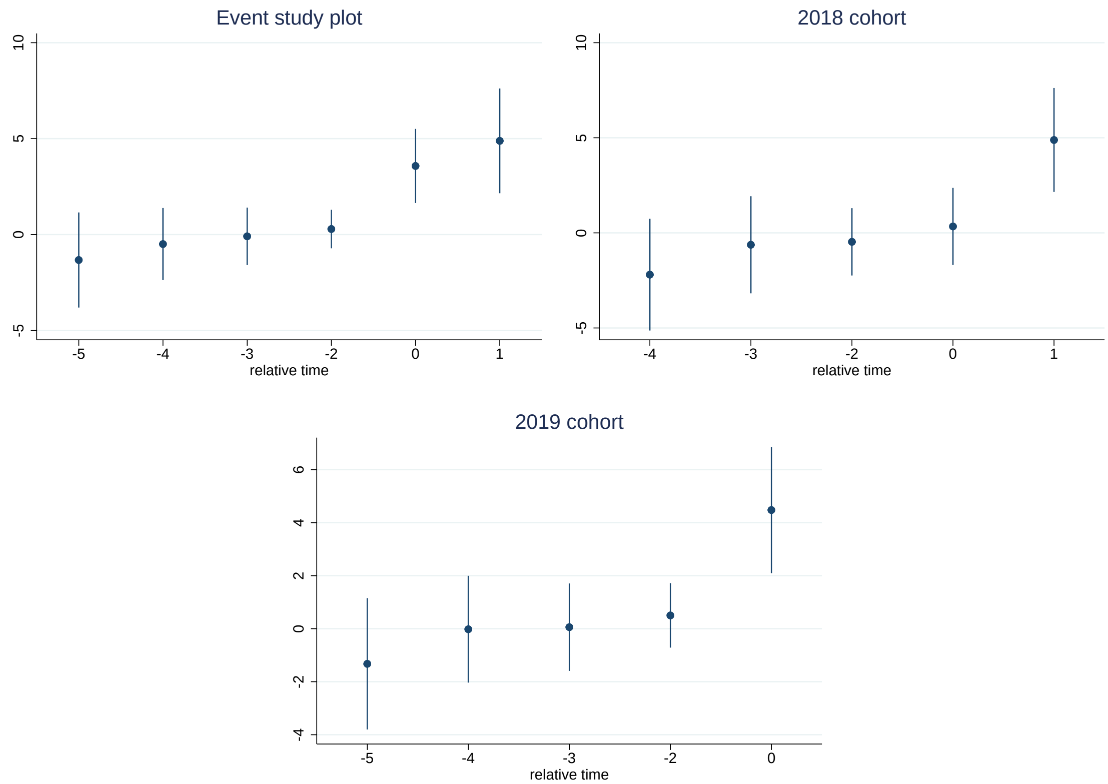
combined effect of these two interventions, in the short term, has likely increased confidence and reduced uncertainty, thereby encouraging entrepreneurship.

Table 1.5: Event Study: Estimates of the Effect of SNAI on the Number of Plants

relative time	TWFE	IW	$CATT_{e,l}$ $\hat{\delta}_{2018,l}$	$\hat{\delta}_{2019,l}$
-5	7.246***	-1.325	-	-1.325
	(1.026)	(1.507)		(1.507)
-4	5.362***	-0.493	-2.195	-0.019
	(0.809)	(1.142)	(1.786)	(1.225)
-3	3.530***	-0.091	-0.626	0.058
	(0.732)	(0.911)	(1.552)	(1.003)
-2	2.136***	0.29	-0.471	0.502
	(0.535)	(0.611)	(1.076)	(0.74)
0	1.418	3.576***	0.339	4.477***
	(0.947)	(1.174)	(1.232)	(1.448)
1	1.604	4.884***	4.884***	-
	(1.604)	(1.661)	(1.661)	
N	1,788	1,788	1,788	1,788

Note: Robust Std. Err. clustered at the municipal level in parentheses. *** p < 0.01, ** p < 0.05, * p < 0.1. TWFE stands for Two Way Fixed Effects; IW stands for interaction-weighted estimator; the column $\hat{\delta}_{2018,l}$ reports the treatment effect on 2018 cohort at different relative times. Similarly for $\hat{\delta}_{201p,l}$. The values in column IW are a weighted mean of $\hat{\delta}_{e,l}$ for $e = 2018, 2019$.

Figure 1.3: Event Study Plot: The Effect of SNAI on the Number of Plants



Note: the upper left picture reports the event study plot of the policy. The upper right figure reports the CATT's for 2018 cohort, whereas the lower left panel reports the CATT's for the 2019 cohort. Vertical bars represent the 95% confidence intervals.

1.6.3 Robustness Checks of Municipal Effects

To corroborate the results presented in the previous section, we report, in Table 1.6, several robustness checks in which we re-estimate the IW model by adding covariates. The results displayed in each column of Table 1.5 are virtually identical and confirm the main result, that is, the policy shows a positive and significant impact on the number of plants at the municipal level.

In order to take into account the skewed distribution of the number of plants at municipality level, we also repeated the estimates by re-weighting the observations by using the inverse-probability weighting (IPW) methodology as well as the standardised IPW which makes the estimator more stable in the presence of values of the propensity score close to zero (Xu et al., 2010). The results are reported in Tables A1 and A2 in the Appendix.

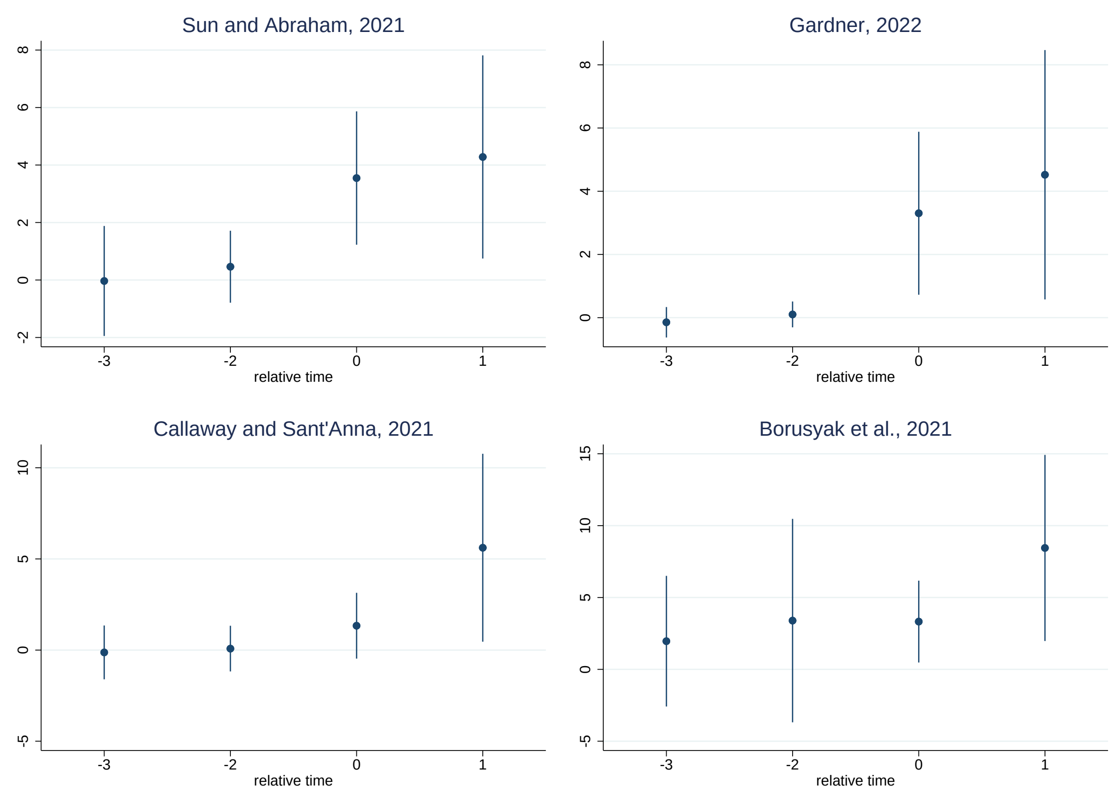
Moving from this result, another check consists of using alternative estimators that recently appeared in the literature of event study analysis. In particular, we have re-estimated column (4) of Table 1.6 using Gardner (2022), Callaway and Sant'Anna (2021) and Borusyak et al. (2021). Evidence is provided in Figure 1.4 which clearly shows that the same result is attainable independently of the estimator. In particular, the policy starts producing effects at relative time 0 and no anticipatory effects are detected. Finally, the same exercise has been repeated without covariates in order to match the column "IW" of Table 1.5 and the result still applies, see Figure A1 in the Appendix.

Table 1.6: Robustness Checks

relative time	(1)	(2)	(3)	(4)
-5	-1.173	-	-	-
	(1.547)			
-4	-0.266	0.165	0.516	0.516
	(1.171)	(1.249)	(1.34)	(1.34)
-3	-0.024	-0.07	-0.033	-0.033
	(0.934)	(0.942)	(0.976)	(0.976)
-2	0.394	0.374	0.463	0.463
	(0.624)	(0.624)	(0.639)	(0.639)
0	3.838***	3.843***	3.547***	3.547***
	(1.197)	(1.19)	(1.183)	(1.183)
1	4.66***	4.659***	4.281**	4.281**
	(1.721)	(1.802)	(1.803)	(1.803)
N	1,746	1,456	1,428	1,428
eco	✓	✓	✓	✓
demo		✓	✓	✓
gov			✓	✓
geo				✓

Note: Robust Std. Err. clustered at municipal level in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. The bottom of the table indicates the sets of covariates included in each column. *eco*=average income of inhabitants; *demo* (log) population; *gov* characteristics of the municipal council; *geo* geographic characteristics.

Figure 1.4: Event Study plot. Four Alternative Estimators



Note: The picture reports the event study plots obtained by applying four different estimators. Sun and Abraham (2021) upper left, Gardner (2022) upper right, Callaway and Sant'Anna (2021) lower left and Borusyak et al. (2021) lower right. Vertical bars represent the 95% confidence intervals.

1.6.4 The Effects of the Policy on the Neighborhood

The results obtained so far show that, in the treated municipalities, the number of plants increased significantly over the first two years. Moving on from this piece of information, it is relevant to ask whether there have been spillover effects on neighbouring municipalities. In principle, positive spillovers may have taken place as more active and thriving municipalities may somehow engender a sort of positive dynamics on the firms located in the neighbourhood. Positive spillovers are especially anticipated for the first type of intervention, which focuses on enhancing essential services. By reducing the distance to these services, municipalities that are within the treatment area but located nearby may also benefit. In a similar way, one can argue that firms may have relocated their productive activity to neighbouring treated municipalities, to benefit from the treatment. This effect is more closely related to the second type of intervention, where business incentives are limited to the area within the treated municipality. Under this last scenario, the policy would generate negative spillovers. Both views are legitimate and the final verdict lies only on empirical grounds.

To answer this relevant question, we have followed the approach taken by Kline and Moretti (2014a) and the ensuing literature. Notably, the authors evaluating the effects of the Tennessee Valley Authority, a place-based policy consisting of the construction of dams and transportation canals on a variety of outcomes such as employment in agriculture and manufacturing, raise the problem of potential spillovers. They notice that by keeping in the non-treated sample counties on the other side of the Authority border, the effects were likely underestimated. As a solution, the authors remove counties that share borders with the Authority. However, a potential limitation of this analysis is that we do not know to what extent spillovers are in order, if any. As a solution, we removed units located in rings of increasing radius. To make the strategy as clear as possible, let us suppose a situation in which spillovers are present only up to d kilometers from the treated units and we run the estimate removing units located once up to d km and once up to f km. The two estimates are expected to produce similar results because, in both cases, the spillovers have been removed, not being present in the ring of form:

$$f - d \quad \text{with } f > d \tag{1.7}$$

As a corollary, and also as a check on that claim, by repeating the estimates without the units in the $f - d$ ring, we expect the estimates to be close to the one obtained without discarding any unit from the sample because we are still keeping units affected by spillovers among the non-treated units, i.e, in the radius d. Table 1.7 implements this strategy. For ease of reference and to see how and if the estimates change, Column (4) of Table 1.5 is reported in Column (1) of Table 1.7. It represents the case in which no potential spillovers are detected. Column (2) of Table 1.7 reports the estimates obtained by discarding neighbours located within 10km from the treated units. Similarly in Column (3), where the radius is set at 20km. Finally, Column (4) removes from the sample the units located in the ring $10 - 20$km from the treated municipalities. In the notation used above $d = 10$, $f = 20$.

Table 1.7: Event Study. Estimates of the Effect of SNAI on the Number of Plants Removing Neighbours

relative time	(1) 0km	(2) 10km	(3) 20km	(4) 10-20km
-4	0.516	0.432	0.515	0.595
	(1.34)	(1.38)	(1.38)	(1.34)
-3	-0.033	-0.017	0.05	0.033
	(0.976)	(0.999)	(1.002)	(0.979)
-2	0.463	0.425	0.495	0.528
	(0.639)	(0.661)	(0.662)	(0.64)
0	3.547***	7.991***	8.017***	3.415***
	(1.183)	(1.63)	(1.719)	(1.218)
1	4.281***	9.998***	9.959***	4.034**
	(1.803)	(2.273)	(2.387)	(1.842)
N	1,428	1,378	1,337	1,423
eco	✓	✓	✓	✓
demo	✓	✓	✓	✓
gov	✓	✓	✓	✓
geo	✓	✓	✓	✓

Note: Robust Std. Err. clustered at the municipal level in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. The bottom of the table indicates the sets of covariates included in each column. *eco*=average income of inhabitants; *demo* (log) population; *gov* characteristics of the municipal council; *geo* geographic characteristics. The estimates have been carried out by using the interacted-weighted (IW) estimators. Column (1) coincides with (4) of Table 1.5, the others have been obtained by removing from the control sample the units located within the radius of 10km, 20km, 10-20km from treated units in column (2), (3) and (4), respectively.

Interestingly, in Column (2) the estimated coefficients at relative time $l = 1$ and $l = 0$ are statistically significant at 1% and higher than the corresponding values in Column (1), where no neighbours were dropped, whereas, no pre-trends are observed. As far as Column (3) is concerned, we do not observe appreciable differences with respect to Column (2). This joint evidence supports the positive spillover hypothesis for units located within 10km from the beneficiaries. In other words, the policy seems to have generated positive effects also on the number of plants of neighbouring municipalities located within a radius of 10km. Eventually, Column (4) confirms the result just claimed that the estimated coefficients are close to those in Column (1), where no units were dropped. These results are in line with Kline and Moretti (2014a).

Having not found any effect on the treated units in terms of the net immigration flow, it is implausible to expect spillovers on neighbours. However, for the sake of completeness, the exercise has been repeated also on these outcomes finding no effects. Evidence is given in Table A3 in the Appendix.

1.7 Conclusions

This study examines the impact of financial assistance provided to Inner Areas in addressing depopulation and promoting business growth in the municipalities that received support. To achieve this, we utilise a particular place-based policy known as the SNAI, which has been implemented in Italy. While the overall policy has limitations, such as the relatively small number of municipalities that received treatment and the limited funding available, it is worth noting that these small, treated municipalities can experience a noticeable positive economic effect due to the targeted financial support. This unique policy implementation provided us with the opportunity to estimate its effects on these communities. For this purpose, we used a detailed panel dataset containing information on Italian municipalities in the years relating to the European programming 2014-2020, and the treated municipalities located in the pilot areas. From a methodological perspective, we have taken advantage of one of the most recent developments in the econometrics of policy evaluation: the IW estimator developed by Sun and Abraham (2021), which generalizes the DiD estimator accounting for staggered entry into treatment. In a non-synchronised treatment setting, the IW estimator uses the last treated as a comparison group, greatly decreasing unobservable heterogeneity, as the comparison group is treated itself. In addition, the causal effects are aggregated by using sensible weightings to obtain an event study analysis.

The results show that the funding did not significantly affect the population structure, but it generated extra plants in the treated municipalities over the first two years. An additional issue tackled in the empirical application is the question of whether the funding has generated spillovers over the neighbouring municipalities, thus reinforcing (in the case of positive spillovers) or crowding out (in the case of negative spillovers) the result obtained. To answer this question an empirical solution has been implemented by following the approach of Kline and Moretti (2014a) who suggest discarding potentially affected units, by concentric rings. Results show that positive spillover effects were occurring, hence corroborating the positive effect of the funding. A caveat in the interpretation of the results is in order. Since the results pertain to a counterfactual situation, the extra plants must not necessarily be construed in terms of new business sites that begin operating. Indeed, they could be plants that would have shut down without treatment.

The policy has not been fully implemented as initially planned, and over the first years, only a small part of the funds was actually invested (notably less than 10 percent of the total funds' programmed amount at the end of 2020). In this regard, the current institutional framework, on which the policy is implemented, creates a mediation mechanism within which the heterogeneity in the quality of the local institution plays a crucial role. The complexity of the decision-making process combined with the reduced administrative capacity of some municipalities could compromise the effectiveness of the SNAI. Nevertheless, this study shows that the policy is promising in the treated municipalities but, given its primary objective of fighting depopulation, it would be desirable to repeat a similar analysis in the future, once the policy is fully implemented, in the new programming cycle and after a certain amount of time, in order to test possible effects on other crucial outcomes such as population, employee demographics and income, if this data becomes available. Of course, the number of plants can only partially reveal the trend of economic activity at the local level, but the limitations of empirical exercises should be considered in the light of the many benefits of having timely empirical evidence on the effectiveness of a programme.

Appendix A

A1. Other Statistics and Econometric Estimation

Table A1: Event study. IPW Estimates of the Effect of SNAI on the Number of Plants.

relative time	(1)	(2)	(3)	(4)	(5)
-4	-0.383	-0.422	-0.705	-0.523	-0.523
	(1.63)	(1.637)	(1.616)	(1.704)	(1.704)
-3	-0.264	-0.469	-0.609	-0.528	-0.528
	(1.139)	(1.168)	(1.148)	(1.181)	(1.181)
-2	0.336	0.293	0.176	0.177	0.177
	(0.719)	(0.735)	(0.726)	(0.739)	(0.739)
0	2.364*	2.354*	2.552*	2.656**	2.656**
	(1.341)	(1.335)	(1.346)	(1.227)	(1.227)
1	3.071**	3.18**	2.836*	2.888*	2.888*
	(1.518)	(1.556)	(1.564)	(1.562)	(1.562)
N	1428	1428	1428	1428	1428
eco		✓	✓	✓	✓
demo			✓	✓	✓
gov				✓	✓
geo					✓

Note: Robust Std. Err. clustered at the municipal level in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. The bottom of the table indicates the sets of covariates included in each column. *eco*=average income of inhabitants; *demo* (log) population; *gov* characteristics of the municipal council; *geo* geographic characteristics. The estimates have been carried out by applying the IPW scheme to the Sun and Abraham (2021) estimator.

Table A2: Event study. Standardized IPW Estimates of the Effect of SNAI on the Number of Plants.

relative time	(1)	(2)	(3)	(4)	(5)
-4	-0.383	-0.413	-0.676	-0.575	-0.575
	(1.63)	(1.637)	(1.627)	(1.724)	(1.724)
-3	-0.264	-0.421	-0.551	-0.517	-0.517
	(1.139)	(1.175)	(1.163)	(1.196)	(1.196)
-2	0.336	0.303	0.194	0.179	0.179
	(0.719)	(0.733)	(0.73)	(0.745)	(0.745)
0	2.364*	2.356*	2.541*	2.695**	2.695**
	(1.341)	(1.336)	(1.353)	(1.214)	(1.214)
1	3.071**	3.154**	2.834*	2.862*	2.862*
	(1.518)	(1.546)	(1.55)	(1.561)	(1.561)
N	1428	1428	1428	1428	1428
eco		✓	✓	✓	✓
demo			✓	✓	✓
gov				✓	✓
geo					✓

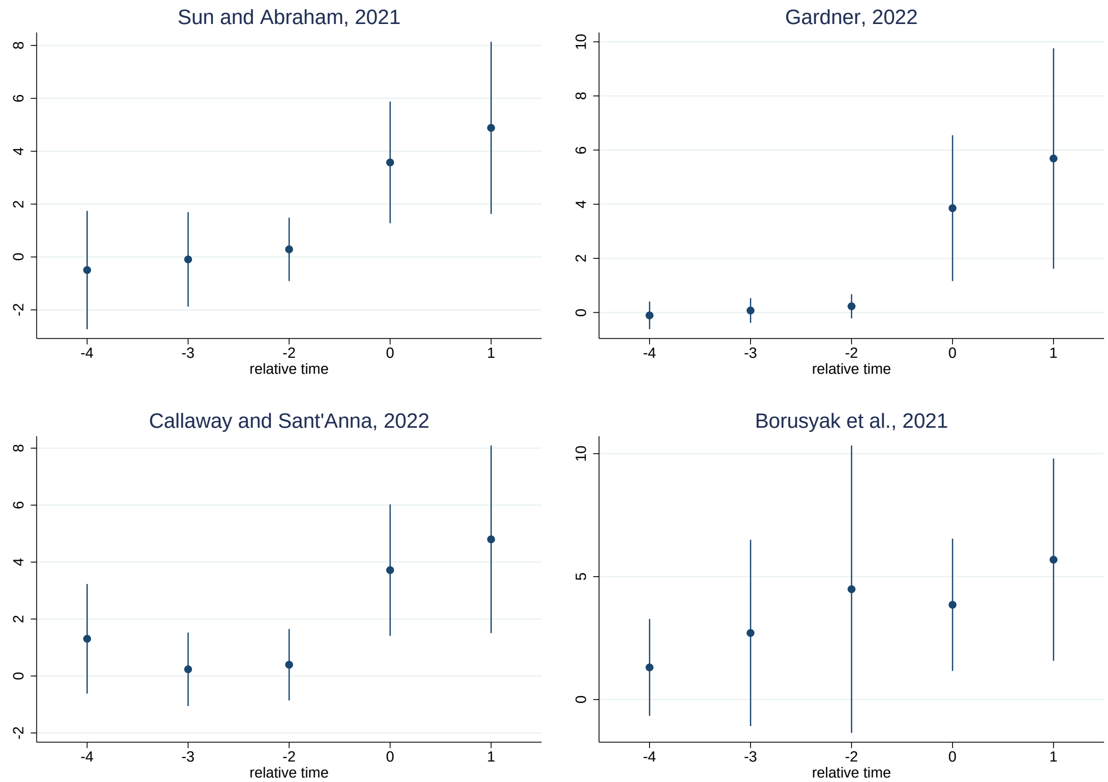
Note: Robust Std. Err. clustered at the municipal level in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. The bottom of the table indicates the sets of covariates included in each column. *eco*=average income of inhabitants; *demo* (log) population; *gov* characteristics of the municipal council; *geo* geographic characteristics. The estimates have been carried out by applying the standardized IPW scheme to the Sun and Abraham (2021) estimator.

Table A3: Event Study. Estimates of the Effect of SNAI on Net Immigration Flow

relative time	(1) 0km	(2) 10km	(3) 20km	(4) 10-20km
-4	1.445	1.512	1.381	1.321
	(3.017)	(3.059)	(3.068)	(3.026)
-3	0.569	0.504	0.404	0.473
	(1.92)	(1.949)	(1.956)	(1.927)
-2	0.189	0.127	0.038	0.105
	(2.066)	(2.097)	(2.105)	(2.073)
0	-1.917	-3.135	-3.048	-1.847
	(1.774)	(2.249)	(2.301)	(1.784)
1	2.953	1.738	1.903	3.1
	(3.955)	(4.298)	(4.336)	(3.957)
N	1,428	1,378	1,373	1,423
eco	✓	✓	✓	✓
demo	✓	✓	✓	✓
gov	✓	✓	✓	✓
geo	✓	✓	✓	✓

Note: Robust Std. Err. clustered at the municipal level in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. The bottom of the table indicates the sets of covariates included in each column. *eco*=average income of inhabitants; *demo* (log) population; *gov* characteristics of the municipal council; *geo* geographic characteristics. The estimates have been carried out by using the interacted-weighted (IW) estimators. Column (1) coincides with (5) of Table 1.4, the others have been obtained by removing from the control sample the units located within the radius of 10km, 20km, 10-20km from treated units in column (2), (3) and (4), respectively.

Figure A1: Event Study Plot. Four Alternative Estimators - Baseline Model with no Covariates



Note: the picture reports the event study plots obtained by applying four different estimators. Sun and Abraham (2021) upper left, Gardner (2022) upper right, Callaway and Sant'Anna (2021) lower left and Borusyak et al. (2021) lower right. The estimates have been obtained without covariates inclusion. Vertical bars represent 95% confidence intervals.

Table A4: Indicators used for Selection of the Pilot Areas

	Indicator	Source
A) Main Characteristics		
a.1	Number of municipalities	Processing by Department for Cohesion Policies
a.2	Number of municipalities in Inner Areas	Processing by Department for Cohesion Policies
a.3	Number of municipalities in Peripheral and Ultra-peripheral Areas	Processing by Department for Cohesion Policies
a.4	Resident population in 2011	Istat - Population and Housing Census 2011
a.5	Number of municipalities in Inner Areas	Istat - Population and Housing Census 2011
a.6	Number of municipalities in Peripheral and Ultra-peripheral Areas	Istat - Population and Housing Census 2011
a.7	Number of municipalities in Inner Areas %	Istat - Population and Housing Census 2011
a.8	Number of municipalities in Peripheral and Ultra-peripheral Areas %	Istat - Population and Housing Census 2011
a.9	Total area in km^2	Istat - Population and Housing Census 2011
a.10	Density per km^2	Istat - Population and Housing Census 2011
B) Demographics		
b.1	Percentage of population aged 0-16 in 2011	Istat - Population and Housing Census 2011
b.2	Percentage of population aged 17-34 in 2011	Istat - Population and Housing Census 2011
b.3	Percentage of resident population aged 65 and over in 2011	Istat - Population and Housing Census 2011
b.4	Percentage of foreign residents in 2011	Istat - Population and Housing Census 2011
b.5	Percentage change of the total population between 1971 and 2011	Istat - Population and Housing Census 2011
b.6	Percentage change of the total population between 2001 and 2011	Istat - Population and Housing Census 2011
b.7	Percentage change of the foreign resident population between 2001 and 2011	Istat - Population and Housing Census 2011
C) Sectoral Specialization and Agriculture		
c.1	Percentage of Utilized Agricultural Area (UAA) in 2010	Istat - Agricultural Census 2000 and 2010
c.2	Percentage change of the Utilized Agricultural Area (UAA) between 1982 and 2010	Istat - Agricultural Census 1982 and 2010
c.3	Percentage change of the Utilized Agricultural Area (UAA) between 2000 and 2010	Istat - Agricultural Census 2000 and 2010
c.4	Percentage of agricultural operators aged up to 39	Istat - Agricultural Census 2010
c.5	Percentage change in the number of agricultural operators aged up to 39 between 2000 and 2010	Istat - Agricultural Census 2000 and 2010
c.6	Incidence of part-time among agricultural operators	Istat - Agricultural Census 2010
c.7	Percentage change in the number of operators with work activities partly carried out on the farm between 2000 and 2010	Istat - Agricultural Census 2000 and 2010
c.8	Percentage of the surface area designated to protected areas	Ministry of Environment, Territory and Sea, 2010
c.9	Percentage of forested surface area	SIAN-INEA - Analysis on Mipaaf data, 2010
c.10 & c.13	Importance index of the agricultural sector in 2001 and 2011.	INEA analysis on Istat data - Agricultural Census, 2001 and 2011
c.11 & c.14	Importance index of the agro-food industry in 2001 and 2011.	INEA analysis on Istat data - Agricultural Census, 2001 and 2011
c.12 & c.15	Importance index of the agro-food sector in 2001 and 2011.	INEA analysis on Istat data - Agricultural Census, 2001 and 2011
c.16	Incidence of companies with PDO and/or PGI productions	Istat – INEA analysis on Agricultural Census data 2010
c.17	Specialization index of Manufacturing Activities	Istat – Active Business Statistical Archive, 2009
c.18	Specialization index of the Energy, Gas, and Water sector	Istat – Active Business Statistical Archive, 2009
c.19	Specialization index of the Construction sector	Istat – Active Business Statistical Archive, 2009
c.20	Specialization index of the Commerce sector	Istat – Active Business Statistical Archive, 2009
c.21	Specialization index of the Other services sector	Istat – Active Business Statistical Archive, 2009
c.22	Number of enterprises per 1000 inhabitants	Business Registry, 2013
c.23	Growth rate of the enterprise stock per 100 enterprises	Business Registry, 2012 and 2013
c.24	Percentage of foreign enterprises	Business Registry, 2013
D) Digital Divide		
d.1	Percentage of population reached by fixed broadband not less than 2 Mbps but less than 20 Mbps (effective capacity)	Ministry of Economic Development, 2013
d.2	Percentage of population reached by fixed broadband not less than 20 Mbps (effective capacity)	Ministry of Economic Development, 2013
d.3	Fixed network digital divide	Ministry of Economic Development, 2013
d.4	Fixed and mobile network digital divide	Ministry of Economic Development, 2013
E) Cultural Heritage and Tourism		
e.1	Number of state and non-state cultural sites	MIBACT, 2012
e.2	Number of inaccessible state and non-state cultural sites	MIBACT, 2012
e.3	Number of visitors	MIBACT, 2012
e.4	% of paying visitors	MIBACT, 2012
e.5	Number of visitors per 1000 inhabitants	MIBACT, 2012
F) Health		
f.1	Services provided per 1000 residents	Ministry of Health, 2012
f.2	Hospitalization rate	Ministry of Health, 2011-2011
f.3	Hospitalization rate of the population over 75 years old	Ministry of Health, 2011-2011
f.4	Avoidable hospitalization rate	Ministry of Health, 2011-2011
f.5	Percentage of elderly >= 65 years residing treated in Integrated Home Care (IHC)	Ministry of Health, 2012
f.6	Percentage of births where the first visit is made from the twelfth week of pregnancy	Ministry of Health, 2011
f.7	Target value of the 75th percentile	Ministry of Health, 2012
f.8	Average number of patients per general practitioner (national orientation: maximum indicated for general practitioners = 1,500)	Managed by the Region
f.9	Average number of patients per freely chosen pediatrician (national orientation: maximum indicated for pediatricians = 800)	Managed by the Region
G) Accessibility		
g.1	Average distance in minutes from non-pole municipalities to the nearest pole	Department for Cohesion Policies
g.2	Average distance in minutes from non-pole municipalities to the nearest pole weighted by population	Department for Cohesion Policies
g.3	Public local bus transport service offering connecting to the regional capital	Managed by the Region
g.4	Public local bus transport service offering connecting to the local pole	Managed by the Region
g.5	Accessibility to the nearest train station (regional or national services) within a 15-minute drive	Managed by the Region
g.6	Accessibility to the nearest train station (regional or national services) within a 15 to 30-minute drive	Managed by the Region
g.7	Intensity of regional train services in relation to the population that can access the service within a 15-minute travel time	Managed by the Region
g.8	Intensity of regional train services in relation to the population that can access the service within a 30-minute travel time	Managed by the Region
g.9	Accessibility to the nearest highway exit	Managed by the Region
g.10	Accessibility to the nearest highway exit	Managed by the Region
g.11	Accessibility to the nearest airport	Managed by the Region
g.12	Accessibility to the nearest port	Managed by the Region
g.13	Synthetic road accessibility indicator of Local Labor Systems where municipalities in the areas are located	ISFORT processing on Istat data 2001.
H) School		
h.1	Average number of schools per institute	Ministry of Education, school year 2013-2014
h.2 & h.15 & h.27	Number of schools	Ministry of Education, school year 2013-2014
h.3	% of municipalities equipped with a primary school	Ministry of Education, school year 2013-2014
h.4 & h.17 & h.29	Average number of students per school (building)	Ministry of Education, school year 2013-2014
h.5 & h.18 & h.30	h.5 & h.18 & h.30- % of students with non-Italian citizenship	Ministry of Education, school year 2013-2014
h.6 & h.19	Disabled students/support teachers	Ministry of Education, school year 2013-2014
h.7 & h.20 & h.31	% of students residing in the same municipality as the school	Ministry of Education, school year 2013-2014
h.8 & h.21 & h.32	Teacher mobility rate	Ministry of Education, school year 2013-2014
h.9 & h.22	% of classes with up to 15 students	Ministry of Education, school year 2013-2014
h.10	% multi-grade classes	Ministry of Education, school year 2013-2014
h.11 & h.23	% full-time classes	Ministry of Education, school year 2013-2014
h.12 & h.24 & h.33	% temporary teachers	Ministry of Education, school year 2013-2014
h.13-14 & h.25-26 & h.34-35	Invalsi test results: average score	Invalsi, school year 2012-2013
h.13-14 & h.25-26 & h.34-35	Invalsi test results: standard deviation	Invalsi, school year 2012-2013
h.16	% of municipalities equipped with a lower secondary school	Ministry of Education, school year 2013-2014
h.28	% of municipalities equipped with an upper secondary school	Ministry of Education, school year 2013-2014
I) Municipal Association		
i.1	Number of municipalities in a union	ANCI, 2013
i.2	% of municipalities in a union	ANCI, 2013
i.3	Number of municipalities in a convention / consortium	ANCI, 2013
i.4	% of Municipalities included in Area Plans (censused)	Isfol, 2013
i.5	Incidence (%) of municipalities in the region area on the total municipalities included in Area Plans	Isfol, 2013

Table A5: Treatments and Details of the Pilot Areas (2018-2020)

Region	Pilot areas	Total Scheduled Cost	Payments Times			Number Municipalities			Total Scheduled Cost per municipality	Treatment per Municipality			Resident Population			Total Scheduled Cost per Inhabitant	Treatment per Inhabitant		
			2018	2019	2020	2018	2019	2020		2018	2019	2020	2018	2019	2020		2018	2019	2020
Abruzzo	Basso Sangro - Trigno	11279540		116053	867576		32	33	341804.24		3516.76	26290.18		19439	18826	599.15		5.97	46.08
Basilicata	Val Sarmento	49528160			2670821			19	2606745.27			140569.51			29920	1655.35			89.27
Basilicata	Montagna Materana	31853361		1691182	5076917		8	8	3981670.17		211397.75	634614.62		10140	9858	3231.22		166.78	515.00
Campania	Alta Irpinia	26026482		574206	1267530		25	25	1041059.28		22968.22	50701.22		59378	58071	448.18		9.67	21.83
Campania	Vallo di Diano	17703369			79621			15	1180224.6			5308.07			56557	313.02			1.41
Emilia-Romagna	Appennino Emiliano	28764229			371855			7	4109175.58			53122.16			32220	892.74			11.54
Emilia-Romagna	Basso Ferrarese	11560645			6546			7	1651520.71			935.13			54450	212.32			0.12
Friuli Venezia Giulia	Alta Carnia	9555000		965240	1405451		20	20	477750		48261.99	70272.57		19098	18700	510.96		50.54	75.16
Liguria	Beigua e Unione Sol	9016815			300647			8	1127101.88			37580.84			17395	518.36			17.28
Lombardy	Alta Valtellina	19376910	1460000	1076613	1076613	5	5	5	3875382	292000	215322.63	215322.64	18572	18516	18329	1057.17	78.61	58.15	58.74
Lombardy	Valchiavenna	21851860	468107	1580794	3289265	12	12	12	1820988.33	39008.88	131732.81	274105.42	24521	24551	24223	902.11	19.09	64.39	135.79
Marche	Basso Pesarese e Anconetano	6736766			132787			9	748529.56			14754.12			31235	215.68			4.25
Apulia	Monti Dauni	77080000			7427627			29	2657931.03			256125.06			54403	1416.83			136.53
Tuscany	Casentino - Valtiberina	10413862			141773			9	1157095.78			15752.57			20014	520.33			7.08
Umbria	Nord - Est Umbria	11893740			259417			7	1699105.77			37059.51			17205	691.30			15.08
Umbria	Sud - Ovest Orvietano	11978482		25740	1575185		20	20	598924.08		1287.02	78759.26		59851	58964	203.15		0.43	26.71
Aosta Valley	Bassa Valle	16298717	877192	1938529	2522219	23	23	23	708639.87	38138.78	84283.89	109661.68	22891	22422	22446	726.13	38.32	86.46	112.37
Aosta Valley	Grand Paradis	7629353			166587			6	1271558.75			27764.55			2278	3349.15			73.13
Veneto	Spettabile Reggenza	11311458			605536			7	1615922.55			86505.16			20334	556.28			29.78
	Total values	**389858749**	**2805298**	**7968357**	**29243973**	**40**	**145**	**269**	**1449289.03**	**70132.46**	**54954.19**	**108713.66**	**65984**	**233395**	**565428**	**948.39**	**45.34**	**55.30**	**72.48**

Source: Author's processing on Department for Cohesion Policies data.

Note: The data reported represent monetary values for thousands and millions of euros.

Bibliography

Accetturo, A. and De Blasio, G. (2012). Policies for local development: An evaluation of italy's "patti territoriali". *Regional Science and Urban Economics*, 42(1-2):15–26.

Acemoglu, D. (2002). Technical change, inequality, and the labor market. *Journal of economic literature*, 40(1):7–72.

Acemoglu, D. and Restrepo, P. (2020). Robots and jobs: Evidence from us labor markets. *Journal of political economy*, 128(6):2188–2244.

Albanese, G., Barone, G., and De Blasio, G. (2022). Populist voting and losers' discontent: Does redistribution matter? *European Economic Review*, 141:104000.

Albanese, G., Barone, G., and De Blasio, G. (2023). The impact of place-based policies on interpersonal income inequality. *Economica*.

Albanese, G., De Blasio, G., and Locatelli, A. (2021). Does eu regional policy promote local tfp growth? evidence from the italian mezzogiorno. *Papers in Regional Science*, 100(2):327–348.

Alesina, A., Cassidy, T., and Troiano, U. (2019). Old and young politicians. *Economica*, 86(344):689–727.

Alesina, A. and Rodrik, D. (1994). Distributive politics and economic growth. *The quarterly journal of economics*, 109(2):465–490.

Alvaredo, F., Atkinson, A. B., Piketty, T., and Saez, E. (2013). The top 1 percent in international and historical perspective. *Journal of Economic perspectives*, 27(3):3–20.

Andini, M. and De Blasio, G. (2016). Local development that money cannot buy: Italy's contratti di programma. *Journal of Economic Geography*, 16(2):365–393.

Arestis, P., Martin, R., and Tyler, P. (2011). The persistence of inequality? *Cambridge Journal of Regions, Economy and Society*, 4(1):3–11.

Atkinson, A. B. and Piketty, T. (2007). *Top incomes over the twentieth century: a contrast between continental european and english-speaking countries*. oup Oxford.

Austin, B. A., Glaeser, E. L., and Summers, L. H. (2018). Jobs for the heartland: Place-based policies in 21st century america. Technical report, National Bureau of Economic Research.

Barca, F. (2009). An agenda for a reformed cohesion policy. a place-based approach to meeting european union challenges and expectations. Technical report, European Commission. Bruxelles.

Barca, F., Casavola, P., and Lucatelli, S. (2014). Strategia nazionale per le aree interne: definizione, obiettivi, strumenti e governance. *Materiali ULAV*, 31(10).

Barca, F., McCann, P., and Rodríguez-Pose, A. (2012). The case for regional development intervention: place-based versus place-neutral approaches. *Journal of regional science*, 52(1):134–152.

Barone, G. and De Blasio, G. (2023a). Place-based policies in the italian case, part 1: A lot of money for little or no growth. *VoxEU. org*.

Barone, G. and De Blasio, G. (2023b). Place-based policies in the italian case, part 2: Mind the negative side effects. *VoxEU. org*.

Bartik, T. J. (2020). Using place-based jobs policies to help distressed communities. *Journal of Economic Perspectives*, 34(3):99–127.

Basile, R., Centofanti, F., Giallonardo, L., and Licari, F. (2023). Migration responses to earthquakes: Evidence from italy. *Italian Economic Journal*, pages 1–23.

Becker, S. O., Egger, P. H., and Von Ehrlich, M. (2010). Going nuts: The effect of eu structural funds on regional performance. *Journal of Public Economics*, 94(9-10):578–590.

Becker, S. O., Egger, P. H., and Von Ehrlich, M. (2012). Too much of a good thing? on the growth effects of the eu's regional policy. *European Economic Review*, 56(4):648–668.

Becker, S. O., Egger, P. H., and von Ehrlich, M. (2018). Effects of eu regional policy: 1989-2013. *Regional Science and Urban Economics*, 69:143–152.

Blouri, Y. and Ehrlich, M. V. (2020). On the optimal design of place-based policies: A structural evaluation of eu regional transfers. *Journal of international economics*, 125:103319.

Borusyak, K., Jaravel, X., and Spiess, J. (2021). Revisiting event study designs: Robust and efficient estimation. *arXiv preprint arXiv:2108.12419*.

Bourguignon, F. and Morrisson, C. (1998). Inequality and development: the role of dualism. *Journal of development economics*, 57(2):233–257.

Bronzini, R. and De Blasio, G. (2006). Evaluating the impact of investment incentives: The case of italy's law 488/1992. *Journal of urban Economics*, 60(2):327–349.

Brunori, P., a. S. N., Ravagli, L., and Maitino, M. L. (2021). Distant and different?: lockdown and inequalities in italy. *Economia Pubblica*, 2:39–54.

Bucci, V., Ferrara, G., and Resce, G. (2023). Local government spending efficiency and fiscal decentralization: evidence from italian municipalities. *Applied Economics*, pages 1–16.

Busso, M., Gregory, J., and Kline, P. (2013). Assessing the incidence and efficiency of a prominent place based policy. *American Economic Review*, 103(2):897–947.

Callaway, B. and Sant'Anna, P. H. (2021). Difference-in-differences with multiple time periods. *Journal of Econometrics*, 225(2):200–230.

Capello, R., Fratesi, U., and Resmini, L. (2011). *Globalization and regional growth in Europe: Past trends and future scenarios.* Springer Science & Business Media.

Carillo, M. and Jappelli, T. (2022). Pandemics and regional economic growth: evidence from the great influenza in italy. *European Review of Economic History*, 26(1).

Carlino, G. A. (1980). Contrasts in agglomeration: New york and pittsburgh reconsidered. *Urban Studies*, 17(3):343–351.

Cerniglia, F., Saraceno, F., and Wattt, A. (2021). *Great reset. 2021 European public investment outlook*, volume OpenBook Publishers.

Cerqua, A. and Pellegrini, G. (2014). Do subsidies to private capital boost firms' growth? a multiple regression discontinuity design approach. *Journal of Public Economics*, 109:114–126.

Cerqua, A. and Pellegrini, G. (2017). Industrial policy evaluation in the presence of spillovers. *Small Business Economics*, 49:671–686.

Cerqua, A. and Pellegrini, G. (2018). Are we spending too much to grow? the case of structural funds. *Journal of Regional Science*, 58(3):535–563.

Cerqueti, R. and Ausloos, M. (2015). Statistical assessment of regional wealth inequalities: the italian case. *Quality & Quantity*, 49:2307–2323.

Cerqueti, R., Coppier, R., Girardi, A., and Ventura, M. (2022). The sooner the better: lives saved by the lockdown during the covid-19 outbreak. the case of italy. *The Econometrics Journal*, 25(1).

Cerulli, G. et al. (2015). Econometric evaluation of socio-economic programs. *Advanced studies in theoretical and applied econometrics series*, 49:198–199.

Cerulli, G. and Ventura, M. (2019). Estimation of pre-and posttreatment average treatment effects with binary time-varying treatment using stata. *The Stata Journal*, 19(3):551–565.

Chattopadhyay, R. and Duflo, E. (2004). Women as policy makers: Evidence from a randomized policy experiment in india. *Econometrica*, 72(5):1409–1443.

Checchi, D., Jappelli, T., Marino, I., and Scognamiglio, A. (2024). Inequality trends in a slow-growing economy: Italy, 1990–2020. *Fiscal Studies*, 45(3):377–392.

Chen, B., Lu, M., Timmins, C., and Xiang, K. (2019). Spatial misallocation: Evaluating place-based policies using a natural experiment in china. Technical report, National Bureau of Economic Research.

Ciani, E. and De Blasio, G. (2015). European structural funds during the crisis: evidence from southern italy. *IZA Journal of Labor Policy*, 4(1):1–31.

Ciani, E., Grompone, A., and Olivieri, E. (2024). Jobs for the long-term unemployed: Place-based policies in depressed areas. *Italian Economic Journal*, pages 1–42.

Cipolloni, C. (2021). Le politiche di contrasto al fenomeno dello spopolamento nelle aree interne. *Italian Papers on Federalism*, (3):52–79.

Clarke, G. R. (1995). More evidence on income distribution and growth. *Journal of development Economics*, 47(2):403–427.

Combes, P.-P. and Gobillon, L. (2015). The empirics of agglomeration economies. In *Handbook of regional and urban economics*, volume 5, pages 247–348. Elsevier.

Crescenzi, R. and De Filippis, F. (2016). Cohesion policy and rural development. In *Handbook on Cohesion Policy in the EU*, pages 325–338. Edward Elgar Publishing.

Crescenzi, R. and Giua, M. (2020). One or many cohesion policies of the european union? on the differential economic impacts of cohesion policy across member states. *Regional Studies*, 54(1):10–20.

Criscuolo, C., Martin, R., Overman, H. G., and Van Reenen, J. (2019). Some causal effects of an industrial policy. *American Economic Review*, 109(1):48–85.

Cusimano, A., Mazzola, F., and Barde, S. (2021). Place-based policy in southern italy: Evidence from a dose–response approach. *Regional Studies*, 55(8):1442–1458.

De Renzis, A., Faggian, A., and Urso, G. (2022). Distant but vibrant places. local determinants of adaptability to peripherality. *Tijdschrift voor economische en sociale geografie*, 113(5):483–501.

Dipartimento per le politiche di coesione (2016). Relazione annuale sulla strategia nazionale per le aree interne. 2015. Technical report, Ministero per la Coesione Territoriale e il Mezzogiorno.

Dipartimento per le politiche di coesione (2017). Relazione annuale sulla strategia nazionale per le aree interne. 2016. Technical report, Ministero per la Coesione Territoriale e il Mezzogiorno.

Dipartimento per le politiche di coesione (2018). Relazione annuale sulla strategia nazionale per le aree interne. 2017. Technical report, Ministro per la Coesione Territoriale e il Mezzogiorno.

Dipartimento per le politiche di coesione (2019). Relazione annuale sulla strategia nazionale per le aree interne. 2018. Technical report, Ministero per il Sud.

Dipartimento per le politiche di coesione (2020). Relazione annuale sulla strategia nazionale per le aree interne. 2019. Technical report, Ministero per il Sud e la Coesione territoriale.

Dipartimento per le politiche di coesione (2021). Relazione annuale sulla strategia nazionale per le aree interne. 2020. Technical report, Ministero per il Sud e la Coesione territoriale.

Ehrlich, M. v. and Seidel, T. (2018). The persistent effects of place-based policy: Evidence from the west-german zonenrandgebiet. *American Economic Journal: Economic Policy*, 10(4):344–374.

European Commission (2014). Accordo di partenariato 2014-2020. Technical report, European Commission, Bruxelles.

Farole, T., Rodríguez-Pose, A., and Storper, M. (2011). Cohesion policy in the european union: growth, geography, institutions. *JCMS: Journal of Common Market Studies*, 49(5):1089–1111.

Fattorini, L., Ghodsi, M., and Rungi, A. (2020). Cohesion policy meets heterogeneous firms. *JCMS: Journal of Common Market Studies*, 58(4):803–817.

Felice, E. (2018). Regional income inequality in italy in the long run (1871–2010): Patterns and determinants. In *The Economic Development of Europe's Regions*, pages 177–203. Routledge.

Ferrara, A. R., Dijkstra, L., McCann, P., and Nisticó, R. (2022). The response of regional well-being to place-based policy interventions. *Regional Science and Urban Economics*, 97:103830.

Ferrara, A. R., McCann, P., Pellegrini, G., Stelder, D., and Terribile, F. (2017). Assessing the impacts of cohesion policy on eu regions: A non-parametric analysis on interventions promoting research and innovation and transport accessibility. *Papers in Regional Science*, 96(4):817–842.

Fields, G. S. (1979). A welfare economic approach to growth and distribution in the dual economy. *The Quarterly Journal of Economics*, 93(3):325–353.

Funk, P. and Gathmann, C. (2015). Gender gaps in policy making: Evidence from direct democracy in switzerland. *Economic Policy*, 30(81):141–181.

Gallo, G. and Pagliacci, F. (2019). Widening the gap: the influence of 'inner areas' on income inequality in italy. *Economia Politica*, 37:197–221.

Gardner, J. (2022). Two-stage differences in differences. *arXiv preprint arXiv:2207.05943*.

Geys, B. (2017). Political dynasties, electoral institutions and politicians' human capital. *The Economic Journal*, 127(605):F474–F494.

Gibbons, S., Lyytikäinen, T., Overman, H., and Sanchis-Guarner, R. (2019). New road infrastructure: The effects on firms. *Journal of Urban Economics*, 110:35–50.

Gibbons, S., Overman, H., and Sarvimäki, M. (2021). The local economic impacts of regeneration projects: Evidence from uk's single regeneration budget. *Journal of Urban Economics*, 122:103315.

Giua, M. (2017). Spatial discontinuity for the impact assessment of the eu regional policy: The case of italian objective 1 regions. *Journal of Regional Science*, 57(1):109–131.

Goodhart, D. (2017). *The road to somewhere: The populist revolt and the future of politics*. Oxford University Press.

Greenstone, M., Hornbeck, R., and Moretti, E. (2010). Identifying agglomeration spillovers: Evidence from winners and losers of large plant openings. *Journal of political economy*, 118(3):536–598.

Guzzardi, D. and Morelli, S. (2024). A new geography of inequality: Top incomes in italian regions and inner areas. Technical report, Center for Open Science.

Hernández-Nicolás, C. M., Martín-Ugedo, J. F., and Mínguez-Vera, A. (2018). Women mayors and management of spanish councils: An empirical analysis. *Feminist Economics*, 24(1):168–191.

Hesse, M. and Rodrigue, J.-P. (2004). The transport geography of logistics and freight distribution. *Journal of transport geography*, 12(3):171–184.

Iammarino, S., Rodríguez-Pose, A., and Storper, M. (2019). Regional inequality in europe: evidence, theory and policy implications. *Journal of economic geography*, 19(2):273–298.

ISTAT (2021). Rapporto annuale 2021. la situazione del paese. Technical report, Governo Italiano.

ISTAT (2022). La geografia delle aree interne nel 2020: vasti territori tra potenzialità e debolezze. Technical report, ISTAT.

Kline, P. (2010). Place based policies, heterogeneity, and agglomeration. *American Economic Review*, 100(2):383–87.

Kline, P. and Moretti, E. (2014a). Local economic development, agglomeration economies, and the big push: 100 years of evidence from the tennessee valley authority. *The Quarterly journal of economics*, 129(1):275–331.

Kline, P. and Moretti, E. (2014b). People, places, and public policy: Some simple welfare economics of local economic development programs. *Annu. Rev. Econ.*, 6(1):629–662.

Knight, J. B. (1976). Explaining income distribution in less developed countries: A framework and an agenda 1. *Oxford Bulletin of Economics and Statistics*, 38(3):161–177.

Kolko, J. and Neumark, D. (2010). Do some enterprise zones create jobs? *Journal of Policy Analysis and management*, 29(1):5–38.

Koster, H. R. and Van Ommeren, J. (2019). Place-based policies and the housing market. *Review of Economics and Statistics*, 101(3):400–414.

Krugman, P. (1991). Increasing returns and economic geography. *Journal of political economy*, 99(3):483–499.

Liberati, P. (2015). The world distribution of income and its inequality, 1970–2009. *Review of Income and Wealth*, 61(2):248–273.

Liberati, P. and Resce, G. (2022). Regional well-being and its inequality in the oecd member countries. *The Journal of Economic Inequality*.

Lu, Y., Wang, J., and Zhu, L. (2019). Place-based policies, creation, and agglomeration economies: Evidence from china's economic zone program. *American Economic Journal: Economic Policy*, 11(3):325–60.

Maranzano, P., M., N., and Romano, R. (2021). The european industrial challenge and the italian nrrp. *PSL Quarterly Review*, 74(298).

Marshall, A. (1890). *Principles of economics*. Macmillan, London.

Mastronardi, L., Giagnacovo, M., and Romagnoli, L. (2020). Bridging regional gaps: Community-based cooperatives as a tool for italian inner areas resilience. *Land Use Policy*, 99:104979.

Mohl, P. and Hagen, T. (2010). Do eu structural funds promote regional growth? new evidence from various panel data approaches. *Regional Science and Urban Economics*, 40(5):353–365.

Moretti, E. (2024). Place-based policies and geographical inequalities. *Oxford Open Economics*, 3(Supplement_1):i625–i633.

Neumark, D. and Kolko, J. (2010). Do enterprise zones create jobs? evidence from california's enterprise zone program. *Journal of Urban Economics*, 68(1):1–19.

Neumark, D. and Simpson, H. (2015). Place-based policies. In *Handbook of regional and urban economics*, volume 5, pages 1197–1287. Elsevier.

Paba, S. and Solinas, G. (2018). In favour of machines (but not forgetting the workers): some considerations on the fourth industrial revolution. *Working in Digital and Smart Organizations: Legal, Economic and Organizational Perspectives on the Digitalization of Labour Relations*, pages 39–63.

Pellegrini, G., Terribile, F., Tarola, O., Muccigrosso, T., and Busillo, F. (2013). Measuring the effects of european regional policy on economic growth: A regression discontinuity approach. *Papers in Regional Science*, 92(1):217–234.

Persson, T. and Tabellini, G. (1994). ls lnequality harmful for growth. *American Economic Review*, 84(3):600–621.

Piketty, T. et al. (2018). Brahmin left vs merchant right: rising inequality and the changing structure of political conflict. *WID. world Working Paper*, 7.

Robinson, S. (1976). A note on the u hypothesis relating income inequality and economic development. *The American economic review*, 66(3):437–440.

Rodríguez-Pose, A. (2018). The revenge of the places that don't matter (and what to do about it). *Cambridge journal of regions, economy and society*, 11(1):189–209.

Rodríguez-Pose, A. and Fratesi, U. (2004). Between development and social policies: the impact of european structural funds in objective 1 regions. *Regional Studies*, 38(1):97–113.

Rodríguez-Pose, A. and Hardy, D. (2015). Addressing poverty and inequality in the rural economy from a global perspective. *Applied Geography*, 61:11–23.

Roth, J. (2024). Interpreting event-studies from recent difference-in-differences methods. *arXiv preprint arXiv:2401.12309*.

Shucksmith, M., Thomson, K. J., Roberts, D., et al. (2005). *The CAP and the regions: the territorial impact of the Common Agricultural Policy.* CABI publishing.

Sonzogno, G. V., Urso, G., and Faggian, A. (2022). Migration propensity of peripheral youth: insights from italy. *Regional Studies, Regional Science*, 9(1):709–726.

Sun, L. and Abraham, S. (2021). Estimating dynamic treatment effects in event studies with heterogeneous treatment effects. *Journal of Econometrics*, 225(2):175–199.

Vendemmia, B., Pucci, P., and Beria, P. (2021). An institutional periphery in discussion. rethinking the inner areas in italy. *Applied Geography*, 135.

Xu, S., Ross, C., Raebel, M. A., Shetterly, S., Blanchette, C., and Smith, D. (2010). Use of stabilized inverse propensity scores as weights to directly estimate relative risk and its confidence intervals. *Value in Health*, 13(2):273–277.

Zheng, L. and Pan, F. (2024). Does place-based policy increase new firm births? evidence from special economic zones in china. *Applied Geography*, 168:103303.

Zheng, S., Sun, W., Wu, J., and Kahn, M. E. (2017). The birth of edge cities in china: Measuring the effects of industrial parks policy. *Journal of Urban Economics*, 100:80–103.

Chapter 2

Predicting Delays of Infrastructure Cohesion Projects:

A Machine Learning Approach*

Abstract

This paper examines the determinants of delays in cohesion infrastructure projects in Italy. We use project completion dates and advanced Machine Learning (ML) techniques to predict and identify important features of delays. The analysis reveals that specific socioeconomic and institutional variables, such as the efficiency of educational institutions, the administrative capacity of local authorities, the amount of funding allocated, and demographic trends, are significant determinants of project timing. Our study also demonstrates the potential of Machine Learning to improve the effectiveness of public investments by addressing both logistical and administrative obstacles, thus contributing to a faster and more efficient implementation of cohesion policies.

JEL classifications: H77; R58; O18; C55

Keywords: Machine learning; Project Delays; Coeshion Policies

*This paper is co-authored with **Professor Giuseppe Coco** (Department of Economics and Finance, University of Bari, Via Camillo Rosalba, 53, Bari BA, 70124, Italy) and with **Professor Giuliano Resce** (Department of Economics, Via Francesco De Sanctis, 86100 Campobasso CB, and International Inequalities Institute, London School of Economics and Political Science, Houghton Street, London, WC2A 2AE, UK) and is published in the Working Paper version on Economics Statistics Discussion Papers, University of Molise, Department of Economics. http://web.unimol.it/progetti/repec/mol/ecsdps/ESDP25099.pdf.

2.1 Introduction

There exists a widespread agreement among scholars that public investment, with a specific emphasis on infrastructure development, plays a crucial role in fostering economic growth and local development (Aschauer, 1989; Leduc and Wilson, 2013; Calderón et al., 2015; Ramey et al., 2021). In this sense, Chaurey and Le (2022) highlight that infrastructure development and maintenance programmes, especially in vulnerable local contexts marked by socioeconomic challenges, can increase employment and economic activity, encouraging the creation of new businesses. Similarly, Khwaja (2009) and Wong et al. (2017) underline how a more careful and inclusive planning and management of infrastructure projects, supported by participatory governance and the active involvement of local communities in decision making, can improve the quality of projects and ensure lasting economic benefits. Acknowledging this, a substantial proportion of cohesion policies is committed to fostering infrastructure initiatives, as evidenced by the allocation of almost half of the European Union's cohesion policy investments during the 2014–2020 programming period towards infrastructure projects (European Commission, 2022). These projects, in the operational implementation phases, are often marked by delays and inefficiencies, due to absorptive capacity constraints, lack of skills, and management issues, that can undermine the benefits of public investment (Espinoza and Presbitero, 2022). Particularly in the cohesion context, local government may lack the capacity to select and manage projects efficiently, leading to cost inflation and delays (Surubaru, 2017; Incaltarau et al., 2020). It has been shown that delays negatively impact project outcomes, with a positive correlation between time and cost overruns (Espinoza and Presbitero, 2022).

This paper investigates the determinants of project delays by leveraging data on infrastructure cohesion projects implemented in Italy in the 2007-2013 wave. We use a Machine Learning (ML) model to predict projects that fail to be completed on schedule. The exploration of factors contributing to project delays has been relatively overlooked in the existing literature, with only a recent exception (Espinoza and Presbitero, 2022). Although traditional statistical and econometric tools have historically probed this issue in alignment with standard economic analysis, contemporary economic literature suggests a shift toward exploiting advancements in ML for addressing policy-related challenges. This perspective contends that the potential of ML surpasses that of conventional econometric techniques for the prediction of policy problems, as argued by Kleinberg et al. (2015). ML techniques are gaining traction in addressing various challenges related to poverty targeting (Jean et al., 2016), evaluating the efficacy of public programmes and spending (Andini et al., 2018), and identifying instances of corruption and political connections (De Blasio et al., 2022; Titl et al., 2024). In the specific context of Italy, recent studies have harnessed ML's potential to forecast the potential bankruptcy of local governments (Antulov-Fantulin et al., 2021), to predict vaccine hesitancy in municipalities (Carrieri et al., 2021), to estimate local mortality and inequality during the COVID-19 pandemic (Cerqua et al., 2021; Cerqua and Letta, 2022), to foresee Geographical Indications areas (Resce and Vaquero-Piñeiro, 2022), and to anticipate dropout rates in higher education (Delogu et al., 2024).

Predicting infrastructure projects which are likely to experience delays could become a valuable tool for proactively addressing potential project weaknesses and evaluating the strengths and weaknesses of different institutional and financial arrangements. Our results indicate that it is possible to forecast which projects will not be completed on time, and that territorial and socioeconomic factors play a significant role.

Italy offers a unique opportunity to investigating delays in cohesion infrastructure projects. Cohesion budgets (125 bil. Euros in the cycle 2007-13) and the number of projects are both very large. Moreover, the Italian Cohesion Agency provides a unique, publicly available database (Opencoesione), with complete records of programmed and actual dates of realisation of all projects financed with cohesion funds. The record also includes financial and sectoral features for each project and geographical information (location). We use this database as a basis, adding information gathered from many different sources for demographic, administrative, economic, geographical, and quality of government indicators, based on the specific location of the project.

The remainder of the paper is structured as follows: Section 2.2 introduces the institutional context in which cohesion programmes are inserted, discussing the structures and rules that regulate interactions between local authorities and supranational institutions along with the evidence about their effectiveness, in term of delays; Section 2.3 presents the dataset used for the analysis, the variable selection criteria, and descriptive statistics that highlight the main characteristics of the data; Section 2.4 describes the empirical approach adopted for data analysis, including statistical models and methodologies for evaluating influential variables; Section 4.4 reports the results of the prediction estimates in term of classification and regression; Section 2.6 shows the territorial heterogeneities that influence delays; Section 4.7 concludes the paper, summarising the main findings, implications for policymakers and future directions for research in this area.

2.2 Institutional Framework and Delayed Projects

A cohesion policy has been one of the original pillars of the European Community, and has continued to be a strategy of the European Union (EU) since its inception in 1957. Territorial disparities have been at the centre of the attention of the founding fathers of the Community. However, only as late as 1975 did it become more relevant in terms of budget and strategy with the creation of the European Regional Development Fund (ERDF), in addition to the existing European Social Fund (ESF). These funds, however, were allotted on the basis of projects proposed on an occasional basis by member states, rather than being constituents of established programmes. The current structure for the EU Regional Policy emerged in the eighties, partly in conjunction with the Single Market pillar. The building of the Single Market was deemed as creating some advantages for peripheral countries, but also as carrying some potential risks of divergence. For these reasons, the enactment of the Single Market and a significant enlargement in 1988, called for a reinforcement of the cohesion policy. The budget for cohesion was almost doubled in the programming period 1989-93 and the relative importance of Cohesion Policy increased considerably. The structure of the intervention was very similar in the subsequent programming cycles. Roughly one-third of the EU Budget is devoted to Cohesion, and the intervention is, for the most part, distributed in two already existing funds, the ERDF and the ESF. Since 2000, Programming periods have extended for 7 years, and the allocation of funds to member states varies according to parameters that change in every programming period. The enlargement process, which brought in new Member States at different levels of development, is the main factor affecting the distribution of funds among nations across time. New entrants usually get a lion's share of cohesion funds.

This notwithstanding, Italy has always been a large beneficiary of the cohesion policy, due to the persistent problem of the *Mezzogiorno*, the largest less-developed area in Western Europe. Cohesion policies are established on an allocation of funds based on the level of development of different regions. All regions get some allocation of funds, but those with GDP per capita below 75% of the EU average get the largest share of funds. In the programming period 2007-13, notwithstanding the accession of most of the new members from Eastern Europe, the share of cohesion funds for Italy was considerable.

Cohesion policies in each country are programmed with a Strategic National Framework (SNF) (Ministero per lo Sviluppo Economico, 2007). The SNF for the period 2007-13 was approved for Italy by the EU Commission and was allocated 28.7 bil. Euro of EU funds (ERDF+ESF). The SNF required Italy to complement those funds with additional funds (co-financing) for 31.6 bil. Euros. The Framework also envisaged the use of an additional national fund, a form of more flexible financing in which the member state had greater discretion about its use - the Underused Areas Fund (in Italian "Fondo Aree Sottoutilizzate", FAS)[1] endowed with 64.4 bil. Euros. Thus, the overall endowment of cohesion policies in Italy for the seven-year programming period was 124.7 bil. Euros. Of these funds, more than 100 bil. Euros were allocated to the Southern less-developed regions.

The Opencoesione database gathers information about all projects financed under this array of funds, even when partially financed by ordinary policies. The overall programming was streamlined according to the ten priorities of the policy. The five priorities receiving the largest share of the funds were: Network and mobility, Competiveness, Energy and Environment, Research and Innovation, and Human Resources. EU funds, National Co-financing Funds, and FAS were allocated to Operative Plans for Regions (POR)[2] and for Ministries (PON)[3] according to priorities. The overall allocation of funds was, thus, quite remarkable and complex. However, the policy ran into considerable difficulties in addition

[1]Later to be renamed definitively in Italian Fondo Sviluppo e Coesione (FSC).
[2]In Italian Piani Operativi Regionali.
[3]In Italian Piani Operativi Nazionali.

to the financial and subsequent debt crises that gripped the country. Nevertheless, it was mainly because of delays in the implementation of the programmes that the government decided to ask the EU for a reformulation of the Framework. In particular, it asked for the possibility of reducing some national co-financing. Delays in delivering programmes were particularly important at the regional level in the *Mezzogiorno*, and, considering the rules that required automatic de-financing of incomplete programs at the second year after the end of the programming period (the n+2 rule), the government decided to re-plan allowing for less national funding. This de-financing cut the national funding by half, cancelling investments for 13.5 bil. Euro in the remaining years of the programming period. At the same time, new regulations and programmes of assistance were adopted for some administrations that were particularly late in delivering.

Contrary to public perception, Italy has never missed targets on final expenditure agreed with the EU Commission, but it is consistently among the countries in which delays in enactment of cohesion policy are larger and that always delivers most of the expenditure in the last two or three years. This was particularly true in the 2007-13 programming period when a large chunk of expenditure occurred in 2015, the last available year for using the funds. Also in the programming period 2014-20, Italy experienced severe delays in expenditure. As of September 2023, 4 months away from the end of the reporting period, Italy spent only 69% of the overall programmed resources (against an EU average of 87%). Several other elements point to a more general difficulty of the Italian public administration in delivering public investment. Public investment, as a proportion of GDP, has been consistently diminishing since 2009, from an average of 3% of GDP in the first decade of the century to a minimum of 1.9% in 2019.

Although general tests about the effectiveness of the EU cohesion policy signal a moderate positive effect (following the seminal paper by Becker et al. (2010)), territorial heterogeneity seems to be large. Results about the effectiveness of the policy in Italy are rather dismal in comparison (Crescenzi and Giua, 2020). A large volume of literature highlights scarce results of individual policy measures (Accetturo and De Blasio, 2012) or for the overall policy (Barone et al., 2016). Many studies point to the relevance of factors such as territorial capital (Fratesi and Perucca, 2018) and especially quality of government (Accetturo et al., 2014; Rodríguez-Pose and Garcilazo, 2015). Crescenzi and Giua (2020) discuss the composition of thematic expenditure and conclude that most of the positive effect of the Cohesion Policy in Europe is concentrated in Germany and the UK. Expenditure in innovation is particularly important. Expenditure in Italy seems on the contrary concentrated on keeping alive low-productivity firms. Coco and Lagravinese (2021) also highlight the differences in the use of funds by identifying classes of expenditure that are more capital intensive, and find that Italian regions display comparatively less capital intensive types of expenditure. It is particularly interesting in the light of our main result, that Coppola et al. (2018) find that European structural funds allocation seems to be less sensitive to the local context (i.e. quality of government) relative to national funds, probably due to rigidity in the allocation and procedures. From this point of view, they argue that European funds are more effective in pursuing long-term development targets.

Delays in delivering investment projects are clearly correlated with the effectiveness and efficiency of expenditure. Problems arising from delays range from increased costs to obsolescence of projects to loss of potential growth. There is also a presumption that delays are (inversely) correlated with the quality of public investment as both depend on the quality of the public administration, see for example Albanese and De Blasio (2018). Awareness of the problem posed by delays is acute also at the institutional level. In 2018 the Agency for Territorial Cohesion published a report on the times of implementation of Public Works, a detailed statistical analysis of correlation with some variables (Agenzia per la Coesione Territoriale, 2018). An interesting feature of the report was the breakdown of implementation times among phases (project, procurement, works). This Report did not present any causal inference, but it highlighted some obvious and less obvious facts. Large projects took substantially longer times of implementation and were more likely to run late; a large fraction of the times of implementation length was due to so called 'cross-through times', that is times that could not be attributed to one phase or the next. A significant heterogeneity among regions emerged, in which some, but not all, southern regions displayed longer times. Regions seemed to perform better overall than other implementing entities. However, these analyses were simple univariate statistical correlation and could not be used for policy advice.

Notwithstanding its huge relevance, the academic literature on the topic of delays is quite scarce. D'Alpaos et al. (2013) show that delays in infrastructure public works at the realisation stage can be due to opportunistic behaviour on the part of firms and that the probability of being levied a penalty for overrun is the main determinant. They then proceed to show that indeed the differential efficiency on Italian courts can explain, in part, differentials in time overrun for public works among regions. This interesting work focuses however only on the final stage, while the evidence shows that delays can be due to stages related purely to administrative procedures. Gori et al. (2017, 2024) highlight the importance of the experience of the administrative staff of municipalities and controls to explain delays. At the international level, Espinoza and Presbitero (2022) show that country characteristics significantly impact delays, with projects in countries with weaker institutions and during periods of increased public investment facing longer delays.

A small strand of literature focuses instead on the best mechanism to prevent delays in the phase of contract award (Reeves et al., 2017), particularly in the case of public-private partnership (PPP) contracts (Palcic et al., 2022). More recently, He et al. (2023) found that negotiated procedures for contract awards may be more timely and efficient and analysed, using UK data, the organisational features that make the award procedure faster.

2.3 Dataset and Descriptive Statistics

2.3.1 Data and Dataset

This work intends to study the most relevant features of the inefficiencies of cohesion projects and, in particular, to evaluate the features and impacts of these on project delays. To this end, we exploit the information heterogeneity of a very rich database for public use, available on the government portal Opencoesione[4], which reports the details of all the territorial cohesion projects implemented in Italy, relating to the European programming cycles, 2007-2013; 2014-2020; 2021-2027[5]. Specifically, for each individual project, the database reports a large amount of procedural information (for example the type of tender, the implementation sector, the nature and operations of the bodies involved in the various procedural phases, presented in Figure 2.2), financial information (quantity of resources assigned to each project, type and nature of financing, etc.), as well as on the temporal dimension (with the expected and actual start and end dates of the projects and of the multiple intermediate phases). Furthermore, it provides details on the territorial scale of reference, with information on the geolocation of the projects, up to the granular municipal level. Consequently, each Italian municipality can be involved in multiple cohesion projects during the annual programming cycles. These interventions can extend on a larger scale than that of a single administrative unit (for example the construction or modernisation of physical infrastructures), affecting multiple regions, provinces and municipalities (multi-municipal interventions).

Operationally, for the predictive estimates, following our empirical design illustrated in detail in the Section 2.4, we proceed to select only the concluded cohesion projects[6] for the 2007-2013 programming cycle. We therefore exclude projects that are not entirely completed in 2024, in order to avoid analysing projects still in the design and implementation phase. Furthermore, we consider only the infrastructural projects that concern the theme of "construction of public works (interventions and installations)", as these are the projects that have the largest construction sites, the largest resources allocated and therefore, as widely shown in the literature, for example Gondia et al. (2020), Egwim et al. (2021), He et al. (2023), the greatest delays and inefficiencies.

Starting from this battery of data, we created our database, which is made up of almost 19 thousand observations at the project scale and around 200 features. We collect endogenous information relating to individual cohesion projects (project-level variables), linked directly or indirectly to the information available from the database, and variables at the municipal level, coming from multiple official statistical sources[7], of an economic,

[4] The database can be freely reached at this link: `https://opencoesione.gov.it/it/`.

[5] At the time of writing (2024), the 2021-2027 European programming cycle is still being.

[6] A cohesion project is defined as concluded if it presents a liquidated financial progress (higher at 95%), as well as an execution phase completed within the monitoring times.

[7] The data on municipal incomes come from the public database of the Ministry of Economy and Finance on tax declarations. The geographical and demographic data are published by the

socio-demographic, environmental, administrative and political nature. The data on institutional quality are at the provincial level. We control also at the territorial level, with macro-regional and regional dummies. Table B2 in the Appendix provides details of all the characteristics used in our estimates.

Our main outcome variable is the status of the project, represented by a dummy variable that takes value of 0 if a project is completed on time, and a value 1 if the project had any delay compared to the deadline declared by the implementing bodies. This variable derives from calculating the days of delay of each project, which is obtained, following Eq. 8, as the difference between the actual end date and the expected end date of each project:

$$d_j = e_j - p_j \tag{2.1}$$

where:

- d_j represents the delay in days for the j-th project;
- e_j is the actual end date of the j-th project;
- p_j is the expected end date of the j-th.

In addition to the estimates conducted with the dummy variable, as a second outcome variable, we exploit the heterogeneity available from the calculation of the days of delay of each project (see Eq. 8). This variable, of a continuous nature, takes on a positive value when the delay is at least one day, 0 when the project ends exactly on the scheduled date, and negative when there is an advance in the closing of the project. This additional estimate allows us to conduct further analyses that also take into account the specific differences in terms of delays between individual infrastructure interventions, thus providing a deeper understanding of the characteristics that condition these temporal differences.

2.3.2 Descriptive Statistics

Tables 2.1, 2.2, 2.3 and Figure 2.1 show descriptive statistics for the distribution of infrastructure cohesion projects 2007-2013, focusing on delays for each Italian region and province and with respect to the type of implementing body. Overall, these statistics suggest heterogeneous territorial effectiveness in the management of cohesion projects, as they show significant regional and provincial disparities in the rate of delayed projects.

Interestingly, Table 2.1 and Figure 2.1b, show that the proportion of late projects out of the total projects in each territorial area does not reflect the Italian regional disparities between North and South (Lagravinese, 2015; Guzzardi et al., 2023). For example, in

Italian National Institute of Statistics. The data on local administrations are from the Ministry of the Interior. The variables on the municipal elections are obtained by manipulating the data of the Ministry of the Interior, made public on the Eligendo portal. The data on public spending by municipalities for technical offices comes from the data of the Open Municipalities Portal. The institutional quality indicators are those created by (Nifo and Vecchione, 2014).

Friuli-Venezia Giulia, which is a Northern Region, 80.12% of projects are delayed.[8] In contrast, some southern regions such as Abruzzo and Molise, have much lower percentages of late projects: 6.92% and 8.31% respectively. The absence of a regional pattern that follows the dualism in Italian development is confirmed at a macro-regional and national level. The South and Centre have the lowest number of delayed projects out of the total in that area (less than 35%). The Islands display the larger percentages of projects with delays (nearly 55%). The North-West and the North-East (with 42% and 39%) show delays similar to the Italian average (38.18%). In general, such high percentages on a macro-territorial and national scale raise important questions about the efficiency of the Italian administrative structures and the project management methodologies adopted (Accetturo et al., 2014; Rodríguez-Pose and Garcilazo, 2015; Albanese and De Blasio, 2018). The Italian dualism reappears importantly in Figure 2.1a, when we consider the number of projects in each region and province, compared to the total number of Italian projects. In this case, the southern regions have more late projects than the northern regions (Baltrunaite et al., 2023).

Table 2.1: Regional Distribution of Infrastructure Cohesion Projects

Region	Total Coeshion Project	Coeshion Projects delayed	Projects delayed on Total Project (%)
Abruzzo	751	52	6.92
Basilicata	1160	396	34.13
Calabria	2053	663	32.29
Campania	5947	2044	34.37
Emilia-Romagna	263	78	29.65
Friuli-Venezia Giulia	312	250	80.12
Lazio	308	88	28.57
Liguria	879	377	42.88
Lombardy	564	123	21.8
Marche	566	187	33.03
Molise	397	33	8.31
Piedmont	880	494	56.14
Apulia	1985	850	42.82
Sardinia	762	266	34.9
Sicily	3216	1908	59.32
Tuscany	697	249	35.72
Trentino-Alto Adige	376	31	8.24
Umbria	589	175	29.71
Aosta Valley	29	17	58.62
Veneto	428	181	42.29
North-east	1379	540	39.16
North-west	2352	1011	42.98
Center	2160	699	32.36
South	12293	4038	32.85
Islands	3978	2174	54.65
N. Observation	**22162**	**8462**	**38.18**

Source: Authors' elaborations on Opencoesione.

[8]The small total number of projects (308) should be considered in this. In Sicily, however, the total number of projects is much higher, as is the number of late projects (equal to 56.14%).

The Table 2.2 and Figures 2.1c, 2.1d, which explore the total days of advance and delay accumulated in cohesion projects, offer a quantitative and more impactful picture and highlight the regional disparities between North and South on this issue (Baltrunaite et al., 2023; Coco et al., 2023). Many southern and island regions, such as Apulia, Calabria, Campania, and Sicily, display an extremely high net negative balance of days of average delay (over 160 days of average delay in the South and more than 340 in the Islands). These findings point to potential systemic and structural problems in project management, which may include bureaucratic inefficiencies, poor planning and coordination, as well as possible deficiencies in transparency and accountability. On the other hand, some provinces in the North-East and the Center show a more balanced or even positive balance, such as Trentino-Alto Adige with an average of 13 days early, which suggests that the projects were completed faster than expected. Such outcomes may be due to better project phase management, effective resource allocation, and solid support infrastructure, which help prevent delays and promote efficient implementation.

Table 2.3 distinguishes delays in infrastructure cohesion projects among different public authorities: Municipalities, Provinces, Regions and State. A large number of late projects emerges at municipal and regional levels, especially in the South and on the Islands. This result confirms the data above, further indicating that the most significant challenges in managing cohesion projects are found at the levels of government closest to citizens, where the administrative capacity and resources involved may be limited (Rodríguez-Pose and Garcilazo, 2015; Albanese et al., 2024). The inefficiency of local institutions in completing cohesion projects within the expected timescales represents a very relevant point for territorial development, as these institutions, during the 2007-2013 European programming cycle, appear to be those most involved in the phase of project implementation, as can be seen from Figure 2.2. Furthermore, these results suggest that some specific regional systems are more resilient and effective than others, probably due to better coordination, less bureaucracy, or a greater availability of resources for timely implementation of projects.

Further descriptive analyses on the delays of infrastructure cohesion projects, implementing and programming bodies, intervention sectors and on the municipal, provincial, and territorial characteristics, are reported in the Appendix 2.7.

Overall, the descriptive statistics presented provide a complex picture of delays in infrastructure cohesion projects in Italy, highlighting the need for a predictive in-depth study approach, to accurately study the characteristics that influence these inefficiencies. A greater understanding of the causes of delays and effective strategies to mitigate them are essential to improve the implementation of cohesion projects, fundamental for the development and balanced growth of the country.

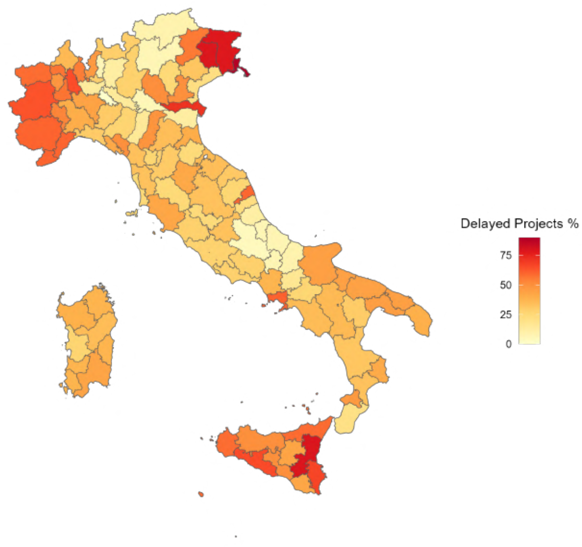
Table 2.2: Days of Project Delay

Region	Total Days Early	Total Days of Delay	Difference between Early and Late	Total Number in Coeshion Project	Average Differential Days by Region
Abruzzo	6488	6253	-235	751	-0.31
Basilicata	18546	161550	143004	1160	123.28
Calabria	27578	595723	568145	2053	276.74
Campania	97751	965901	868150	5947	145.98
Emilia-Romagna	2669	11847	9178	263	34.90
Friuli-Venezia Giulia	7285	69745	62460	312	200.19
Lazio	1458	21444	19986	308	64.89
Liguria	37961	97100	59139	879	67.28
Lombardy	3956	38257	34301	564	60.82
Marche	30497	40473	9976	566	17.63
Molise	3153	7823	4670	397	11.76
Piedmont	9481	177914	168433	880	191.40
Apulia	66553	516917	450364	1985	226.88
Sardinia	33245	56685	23440	762	30.76
Sicily	28653	1368741	1340088	3216	416.69
Tuscany	12666	39011	26345	697	37.80
Trentino-Alto Adige	15385	10497	-4888	376	-13.00
Umbria	38292	44311	6019	589	10.22
Aosta Valley	76	3404	3328	29	114.76
Veneto	10780	27593	16813	428	39.28
North-east	36119	119682	83563	1379	60.60
North-west	51474	316675	265201	2352	112.76
Center	82913	145239	62326	2160	28.85
South	220069	2254167	2034098	12293	165.47
Islands	61898	1425426	1363528	3978	342.77
Total	**452473**	**4261189**	**3808716**	**22437**	**171.85**

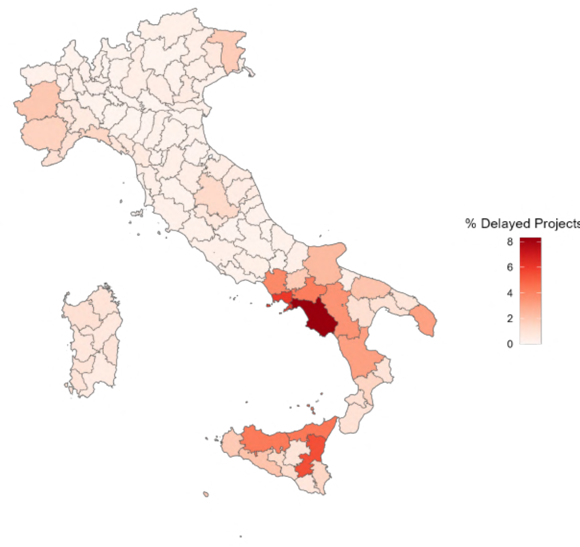
Source: Authors' elaborations on Opencoesione.

Figure 2.1: Provincial Distribution of Delayed Infrastructure Cohesion Projects and Average Days of Delay at Provincial Level

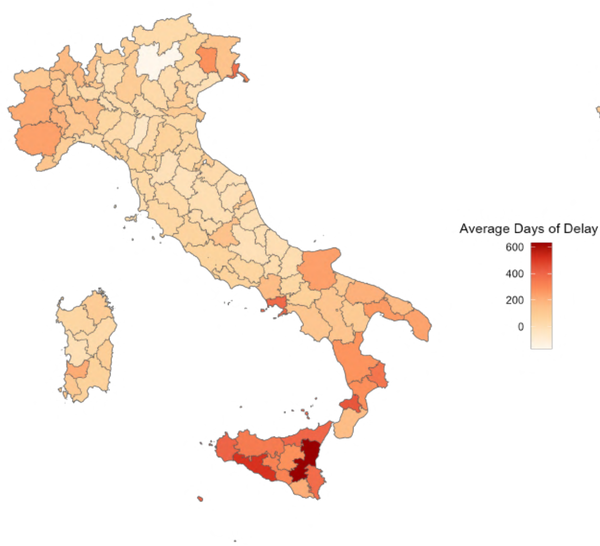
(a) Percentage of Delayed Cohesion Projects per Province on to the Total Cohesion Projects in that Province



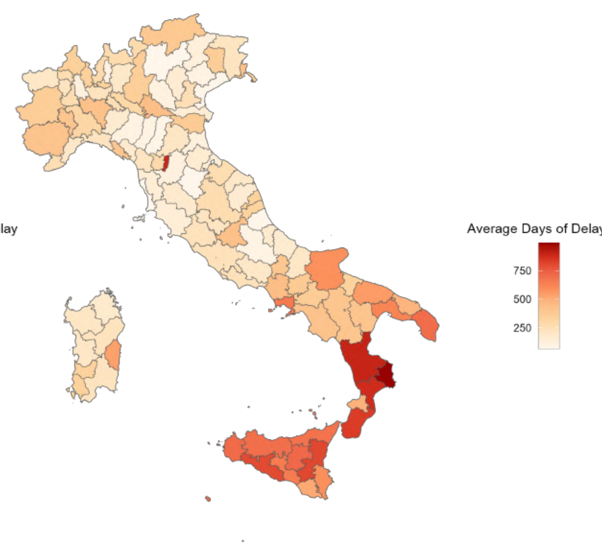
(b) Percentage of Delayed Cohesion Projects per Province on to the Total Cohesion Projects in Italy



(c) Average Days of Delay of Cohesion Projects per Province on the Total Cohesion Projects in that Province



(d) Average Days of Delay of Cohesion Projects per Province Calculated on the Delayed Cohesion Projects in that Province



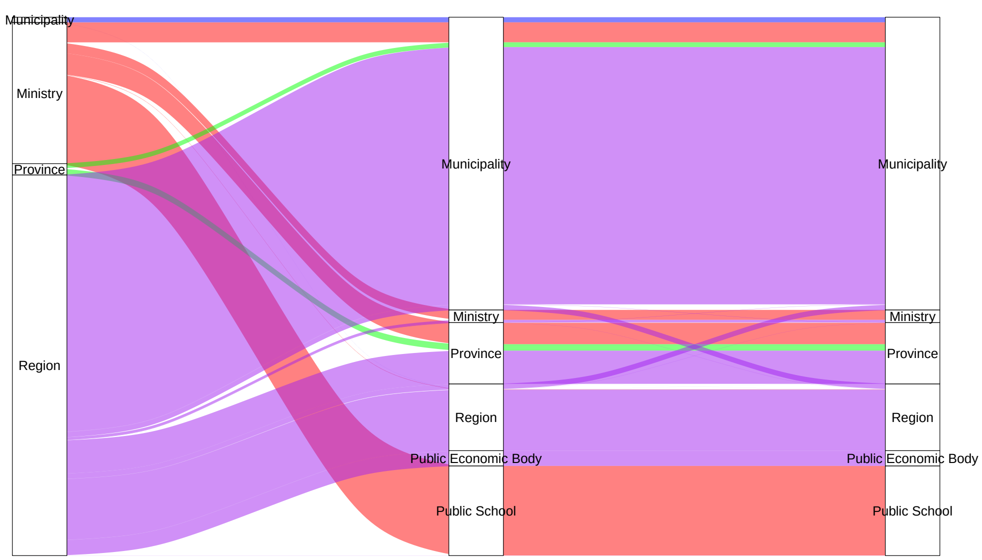
Source: Authors' elaborations on Opencoesione.

Table 2.3: Type of Entities per Macro Area

Entities	Status	Marco Area					Total Status by Entities
		Northeast	North-west	Center	South	Islands	
Municipalies	**0**	262	763	833	3522	581	**5961**
	1	244	507	379	1046	703	**2879**
	Total	**506**	**1270**	**1212**	**4568**	**1284**	**8840**
Province	**0**	197	151	119	461	278	**1206**
	1	147	200	164	158	76	**745**
	Total	**344**	**351**	**283**	**619**	**354**	**1951**
Region	**0**	110	2	14	442	704	**1272**
	1	47	14	4	50	402	**517**
	Total	**157**	**16**	**18**	**492**	**1106**	**1789**
National	**0**	42	123	29	250	34	**478**
	1	3	10	13	171	53	**250**
	Total	**45**	**133**	**42**	**421**	**87**	**728**
Total Status by NUTS	**0**	**611**	**1039**	**995**	**4675**	**1597**	**8917**
	1	**441**	**731**	**560**	**1425**	**1234**	**4391**
	Total	**1052**	**1770**	**1555**	**6100**	**2831**	**13308**

Source: Authors' elaborations on Opencoesione.

Figure 2.2: Flows of Entities in Cohesion Project



Source: Authors' elaborations on Opencoesione.
Note: The first column refers to the programming bodies. The second to the implementing authorities. The third to the beneficiary bodies.

2.4 Empirical Approach

2.4.1 Empirical Design

In our work, the Prediction task is defined following the Eq. 9.

$$\{Features_i\} \xrightarrow{f(\cdot)} Prediction_i \tag{2.2}$$

In particular, for each infrastructural cohesion project i started at the time t, we want to obtain a function $f(\cdot)$ represented by a machine learning model that employs a series of features/predictors known at the time t, able to estimate the probability that an infrastructural cohesion project will have a delay $Delay_{iT}$:

$$\{Pro, SocioEco, Demo, Ter, Geo, IQI, Pol, Tech\}_{it} \xrightarrow{f(\cdot)} Delay_{iT} \tag{2.3}$$

where:

- *Pro*: Project-level features of cohesion projects;
- *SocioEco*: Municipal Socioeconomic features;
- *Demo*: Municipal Demographic features;
- *Ter*: Regional and Macro Area features;
- *Geo*: Municipal Geographical features;
- *IQI*: Provincial Institutional Quality and Government Effectiveness features;
- *Pol*: Municipal Local Policy and Administrative features;
- *Tech*: Municipal Features of Public Spending for Technical Offices;
- *Delay*: Status of delay.

To have a predictive tool useful for policy, we only used features observed at the start date of the project (t) to predict the delay at T, so with $t < T$. In line with standard practices in prediction analyses with supervised machine learning, classification and regression algorithms, conducted in the field of social sciences, see for example Antulov-Fantulin et al. (2021), Bloise et al. (2021), Carrieri et al. (2021), the dataset is randomly divided between the training set and the testing set. In this study, we assign 70% of the data to the training set, while the remaining 30% constitutes the testing set.

As described in Section 2.3, to conduct our ML analyses, we use the dummy of late projects as a dependent variable. Furthermore, to further corroborate our estimates we also implement ML algorithms with a continuous variable. For these analyses, the outcome variable is represented by the days of delay (see Eq. 8), which allows us to overcome data interpretation problems, in consideration of the high heterogeneity present in the days of delay between individual projects. As a robustness analysis, in consideration of the marked territorial and regional divisions present in Italy (Lagravinese, 2015; Guzzardi et al., 2023), especially in reference to the effectiveness of cohesion policies (RodrÍguez-Pose and Fratesi, 2004; Cerqua and Pellegrini, 2018; Crescenzi and Giua, 2020) and the projects implemented with them (Coco et al., 2023), we repeat the main ML estimates, with the binary variable of the projects delay, by splitting our dataset on a macro-regional scale, between North and South[9]. This is in order to detect socio-spatial differences regarding the most important features in predicting the delay of infrastructural cohesion policies.

The optimisation of model parameters is derived from a repeated cross-validation process exclusively on the training set (k-fold = 10, with 5 repetitions).

The accuracy of the model in predicting delays in classification estimates is quantified using the Receiver Operating Characteristics (ROC) curve (Bradley, 1997; Fawcett, 2006), with the test set as the evaluation context. This binary classification delineates late projects as a positive class and early projects as a negative class. In particular, the ROC curve clarifies the competence of the classifier, relating the true positive rate (TPR) to the false positive rate (FPR) at various discrimination thresholds. An uninformative classifier results in an ROC curve that mirrors the line of randomness, producing an area under the curve (AUC) of 0.5, while a flawless classifier boasts an AUC of 1.0. The AUC therefore serves as a benchmark for the effectiveness of the classifier, with higher values indicating superior discriminatory power.

The ROC curve is defined as:

$$TPR(y) = \frac{TP}{TP + FN} \tag{2.4}$$

where:

- TP: is the number of true positives;
- FN: is the number of false negatives.

$$FPR(x) = \frac{FP}{FP + TN} \tag{2.5}$$

where:

- FP: is the number of false positives;
- TN: is the number of true negatives.

In the context of regression, where continuous outcomes such as project delay days are expected, we adopt several performance metrics to evaluate the effectiveness of ML models. We use metrics, such as the Root Mean Square Error (RMSE), the Mean Squared Error (MSE), and the Mean Absolute Error (MAE), which provide a complete view of both the accuracy of the predictions and the fit of the model to the data. Thanks to these indicators, we can effectively compare the performance of the various models and select the most appropriate algorithm for the specific context of use (Botchkarev, 2018; Chicco et al., 2021;

[9]The North with Emilia Romagna, Liguria, Lombardy, Marche, Piedmont, Tuscany, Umbria, Veneto. Friuli Venezia Giulia, Trentino Alto Adige, and Aosta Valley do not present any infrastructure cohesion projects, after the data manipulation operation, for the construction of the dataset. The South with Abruzzo, Apulia, Basilicata, Calabria, Campania, Lazio, Molise, Sicily, Sardinia.

Caravaggio and Resce, 2023).

2.4.2 ML Algorithms

For our estimates, we exploit a battery of eight ML algorithms for classification and regression (discussed below), as well as classic regression models for comparisons (logistic model and linear model)[10].

Elastic Net (EN): is a statistical regression method that combines the penalties of Lasso regression (L1) and Ridge regression (L2), in order to optimise the selection and regularisation of variables in high-dimensional contexts. This algorithm is particularly effective in situations where the variables present multicollinearity, offering a robust compromise between the variable selection capacity of Lasso and the regularisation of Ridge, thus improving the stability and accuracy of the model (Zou and Hastie, 2005; De Mol et al., 2009; Cerulli, 2023). The formulation of Elastic Net is as follows:

$$\min_{\beta} \left\{ \frac{1}{2n} \|\mathbf{y} - \mathbf{X}\beta\|^2 + \lambda \left(\alpha \|\beta\|_1 + \frac{1}{2}(1-\alpha)\|\beta\|_2^2 \right) \right\} \tag{2.6}$$

where:

- β is the vector of coefficients estimating the weights for explanatory variables;
- $\mathbf{y}$ is the vector of observed target values;
- $\mathbf{X}$ is the feature matrix with rows as observations and columns as variables;
- n is the number of observations in the dataset;
- The term $\frac{1}{2n}\|y - X\beta\|^2$ is the mean squared error (MSE), measuring model fit;
- λ is the regularization parameter that controls the intensity of the penalties;
- α is the mixing parameter between Lasso (L1) and Ridge (L2) penalties, with $\alpha = 1$ corresponding to Lasso and $\alpha = 0$ to Ridge;
- $\|\beta\|_1$ and $\|\beta\|_2^2$ are the L1 and L2 norms of β, implementing the Lasso and Ridge penalties respectively, with L1 promoting sparsity and L2 handling multicollinearity.

Support Vector Machine (SVM): Its main objective is to find a hyperplane that optimally separates the different classes of data points. To facilitate this separation, the model transforms the data into a higher dimensional space. This transformation is particularly useful in scenarios where the number of features, and therefore data dimensions, exceeds the number of observations. In these cases, the SVM allows one to effectively manage complexity while avoiding overfitting. It is suitable for applications that require good generalisation ability, even in the presence of a limited number of observations (Hearst et al., 1998; Steinwart and Christmann, 2008). The SVM tries to maximise the margin between classes by minimising $\|\mathbf{w}\|$, while penalising misclassifications with a loss function. The optimal formulation of the SVM in terms of the minimisation problem is:

[10] All ML algorithms and optimisation processes were implemented using R software via the `caret` package (Kuhn, 2015; Kuhn et al., 2020).

$$\min_{\mathbf{w},b} \left\{ \frac{1}{2}\|\mathbf{w}\|^2 + C\sum_{i=1}^{n}\xi_i \right\} \quad \text{with } 1 < i < n \tag{2.7}$$

The formula for the optimal hyperplane in an SVM is as follows:

$$y_i(\mathbf{w}\cdot\mathbf{x}_i - b) \geq 1 - \xi_i \quad \text{with} \quad \xi_i \geq 0 \tag{2.8}$$

where the parameters are defined as:

- $\mathbf{w}$ is the weight vector normal to the hyperplane;
- b is the bias term that shifts the hyperplane;
- C is a regularisation parameter that balances the width of the margin and the amount of margin violations;
- n is the number of training samples;
- ξ_i are the slack variables, which allow for some classification violations for data that are not linearly separable;
- y_i are the class labels (typically +1 or -1);
- $\mathbf{x}_i$ are the input vectors.

Random Forest (RF): is a ML model based on decision trees. Each tree is constructed using bootstrap replication, but unlike Bagging (bootstrap aggregation), RF selects the splitting variable at each node by randomly sampling from the available features. This process helps to "decorrelate" the trees within each bootstrap sample, thereby enhancing exploration of the feature space and improving out-of-sample predictive performance. RF provides a robust and effective solution for complex classification and regression problems, particularly with large datasets (Breiman, 2001; Liu et al., 2012; Cerulli, 2023). The final output of the RF model can be expressed mathematically as follows:

$$\hat{y} = \frac{1}{B}\sum_{b=1}^{B} T_b(x;\Theta_b) \tag{2.9}$$

where:

- $\hat{y}$ is the final predicted value obtained by averaging the predictions of all B trees;
- B is the number of trees in the forest;
- $T_b(x;\Theta_b)$ is the prediction of the b-th decision tree;
- x represents the input features;
- Θ_b denotes the randomly selected parameters used to build the tree b.

Gradient Boosting Machines (GBM): is an algorithm that iteratively improves its predictions, correcting the errors of previous models. This approach is structured as a basic decision tree model which increases its accuracy by adding new models sequentially, resolving the residual errors of the aggregate model up to that point. Each new model in the sequence is then trained to be as effective as possible at correcting the errors left by its predecessors. The algorithm is characterised by hyperparameters such as the learning rate control and the tree depth that are suitably fine-tuned to prevent overfitting. GBM maintains a good balance between adapting to training data and generalising to new data

(Friedman, 2001; Natekin and Knoll, 2013). More specifically, the GBM takes this general form:

$$f(x) = \sum_{b=1}^{B} \beta_b g_b(x, \phi_b) \tag{2.10}$$

where:

- B is the total number of boosting iterations or trees used in the model;
- β_b is the b-th expansion coefficient;
- x represents the input features;
- ϕ_b represents the parameters learned by the b-th model, typically associated with the structure of the decision tree (such as split points, leaf values, etc.).

Decision Tree (DT): This is an ML device based on building trees using binary splitting at each node (Charbuty and Abdulazeez, 2021; Cerulli, 2023). The split is made using the variable and the threshold that minimise statistical criteria such as the Gini index, the error rate, and the entropy. To build an appropriate tree, it is essential to determine the best splits at each node by calculating the proportions p_i of class i instances among the cases in the set S.

The Gini Index G for a set S is calculated as follows:

$$G(S) = 1 - \sum_{i=1}^{c} p_i^2 \tag{2.11}$$

where:

- S is the set of data considered for the split;
- c is the number of classes in set S;
- p_i is the proportion of instances belonging to class i in set S.

Entropy measures the impurity of a node, with higher values indicating more mixed sets. The entropy $H(S)$ for a set S is calculated as:

$$H(S) = -\sum_{i=1}^{c} p_i \log_2 p_i \tag{2.12}$$

where:

- $H(S)$ is the entropy of set S, which measures the impurity of the node;
- p_i is the proportion of instances belonging to class i;
- c is the number of classes in set S.

To determine the best attribute to split the data at each node, the algorithm calculates the Information Gain (IG), which measures the entropy reduction after the split. Information Gain for an attribute A. Is calculated as:

$$IG(S, A) = H(S) - \sum_{t \in T} \frac{|S_t|}{|S|} H(S_t) \tag{2.13}$$

where:

- S is the set of data before the split;
- A represents the attribute for the split;
- T is the set of subsets S_t formed after the split on attribute A;
- $\frac{|S_t|}{|S|}$ is the proportion of instances in subset S_t;
- S_t is the subset of data after the split for branch t;
- $H(S_t)$ is the entropy of subset S_t.

Naive Bayes (NB): This is used only in classification problems when the features used are independent of each other. It estimates the probability that an element belongs to a certain class using the Bayes formula. This model assumes the conditional independence of features given the class label, which simplifies the computation of class probabilities (Rish et al., 2001; Murphy et al., 2006). The probability that an element belongs to a class C_k given features $\mathbf{x} = (x_1, x_2, \ldots, x_n)$ is calculated using Bayes' theorem as follows:

$$P(C_k \mid \mathbf{x}) = \frac{P(C_k)P(\mathbf{x} \mid C_k)}{P(\mathbf{x})} \tag{2.14}$$

where:

- $P(C_k)$ is the prior probability of class C_k;
- $P(\mathbf{x} \mid C_k)$ is the likelihood, which is simplified to $\prod_{i=1}^{n} P(x_i \mid C_k)$ under the independence assumption;
- $P(\mathbf{x})$ is the evidence, often computed as a scaling factor that normalises the probability distribution;
- $\mathbf{x} = (x_1, x_2, \ldots, x_n)$ is the feature vector, where x_i represents the i-th feature;
- C_k is the class label for class k.

Neural Networks (NN): inspired by the functioning of the human brain, they are composed of nodes, or neurons, interconnected and organised in layers. Each node receives input from the previous nodes on which it performs a weighted sum. After calculating the weighted sum, the node transmits the output a to the subsequent nodes. During the learning process, neural networks adapt the weights of the connections through algorithms such as gradient descent, which allow for minimising the error. Neural networks are particularly effective in classification and regression studies in the presence of complex features, thanks to their high learning capacity (Sutskever et al., 2014; Schmidhuber, 2015). This weighted sum is calculated as follows:

$$z = \sum_{i=1}^{n} w_i x_i + b \tag{2.15}$$

where:

- z is the weighted sum (pre-activation value);
- n is the total number of inputs to the neuron;
- w_i represents the weight associated with the input x_i;
- x_i is the i-th input to the neuron;
- b is a bias term.

The output of the node, after applying an activation function σ, is given by:

$$a = \sigma(z) \tag{2.16}$$

where:

- a is the activation (output) of the neuron;
- σ is the activation function, which could be a non-linear function like ReLU, sigmoid, or tanh[11];
- z is the weighted sum (pre-activation value).

K-Nearest Neighbors (KNN): is based on the principle of closeness in similarity, meaning that similar elements tend to be close to each other (Peterson, 2009; Kramer and Kramer, 2013). The algorithm calculates the distance between the query instance and all the instances in the training set, then selects the nearest k instances. The most commonly used distance metrics are Euclidean distance, Manhattan distance, and Minkowski distance.

For classification, the most common class among the k nearest neighbours is assigned to the query instance.

$$\text{Class}(\mathbf{x}_q) = \text{mode}\{\text{Class}(\mathbf{x}_{i1}), \text{Class}(\mathbf{x}_{i2}), \ldots, \text{Class}(\mathbf{x}_{ik})\} \tag{2.17}$$

where:

- $\mathbf{x}_q$ is the query input vector (the point we want to predict);
- k is the number of nearest neighbours;
- $\text{Class}(\mathbf{x}_{i1}), \text{Class}(\mathbf{x}_{i2}), \ldots$ are the classes of the k nearest neighbours.

In the case of regression, instead of assigning the most common class, KNN assigns the average value of the dependent variable of the k nearest neighbours. The formula for regression in KNN is as follows:

$$\text{Value}(\mathbf{x}_q) = \frac{1}{k}\sum_{i=1}^{k} y_i \tag{2.18}$$

where:

- $\mathbf{x}_q$ is the query input vector (the point we want to predict);
- k is the number of nearest neighbours;
- y_i are the values of the dependent variable (target) for the i-th nearest neighbour.

[11] ReLU (Rectified Linear Unit) is commonly used for deep networks due to its efficiency and ability to mitigate the vanishing gradient problem, as it outputs zero for negative values and the input itself for positive values. Sigmoid is often used in binary classification tasks, mapping values between 0 and 1, while tanh maps values between -1 and 1, and is preferred in some cases for hidden layers due to its zero-centred output.

2.5 Results

In this Section we set out our predictive results, i.e. model estimates, focusing on cohesion project delays. We focus on the ability to predict the dependent variables and the key role of the independent variables (features) used in the predictions, described in, described in Section 2.3.1.

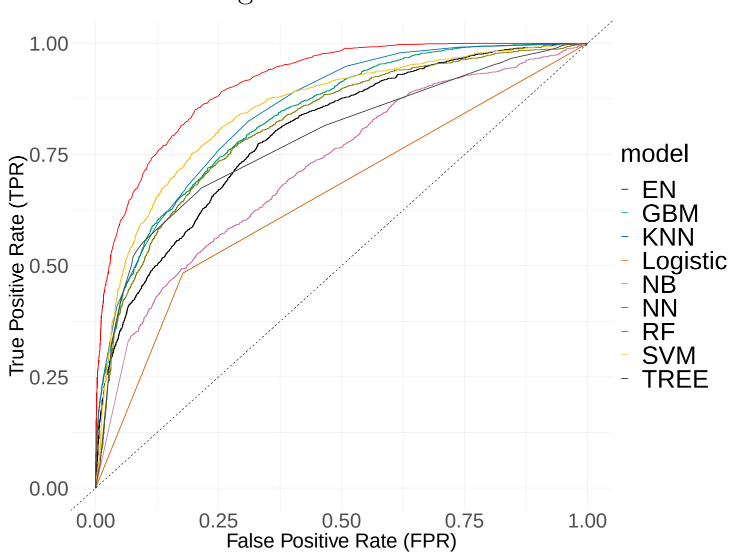
2.5.1 Classification Task

2.5.1.1 Prediction Performance

In this section we evaluate the performance of the various machine learning models on the main outcome variable of our estimates (out-of-sample), represented by the dummy of delayed projects, and subsequently, we evaluate the optimisation of the hyperparameters based on the training data.[12].

The AUC analysis, illustrated in the ROC curve graph (see Figure 2.3), highlights how the Random Forest (RF) model stands out as the best model for performance, compared to other ML algorithms, including regression logistics (Bradley, 1997; Fawcett, 2006). In particular, the AUC shows from a comparative perspective that RF significantly outperforms all other ML models. In practice, this reveals the better performance of RF compared to SVM, GBM, KNN, NB, NN, EN and Logistic Regression, regarding the prediction of delays in infrastructure cohesion projects in Italy. Furthermore, Figure 2.3 shows the heterogeneity among the various models.

Figure 2.3: ROC Curve



[12]For the models analysed on the entire set of Italian infrastructure projects, the best hyperparameters identified are shown in Appendix B1.

The results of the DeLong statistical test (DeLong et al., 1988) show that, for the RF model, the difference between the ROC curve and the remaining models is statistically significant (with p-value < 0.05), in all comparisons between the algorithms. The Z values are also high (ranging from 15.02 to 39.838), since they reflect the robustness of the superiority of the RF model in various comparisons.

These results are in line with previous empirical studies, for example, Antulov-Fantulin et al. (2021), Carrieri et al. (2021), Resce and Vaquero-Piñeiro (2022), which reveal that Tree models are the most predictive algorithms on granular data, especially in classification contexts with binary response variables.

Table 2.4: DeLong's Test

	Z	p-value
RF vs SVM	15.02	0.00
RF vs KNN	16.43	0.00
RF vs GBM	18.32	0.00
RF vs NN	19.31	0.00
RF vs EN	21.64	0.00
RF vs TREE	21.88	0.00
RF vs NB	27.63	0.00
RF vs Logistic	39.83	0.00

Note: Averages calculated across repetitions for 10 different random splits.

In general, the metrics reported in Table 2.5 show high performances of all the estimated algorithms. In particular, the high performance of these models in predicting late projects is confirmed by a No Information Rate lower than the accuracy level (Accuracy > Accuracy Null) and by the other reported metrics (Hand and Till, 2001; Tharwat, 2021). Furthermore, Cohen's Kappa coefficients show values that reflect the trends recorded in the accuracy of the individual algorithms (Cohen, 1960; McHugh, 2012). In fact, the RF emerges as the model with the best agreement (0.615), followed by SVM, Decision Tree, KNN and Neural Network (the latter with a moderate agreement of around 0.5). In contrast, the GBM, EN, NB and logistic regression show lower performances (with an agreement of between 0.45 and 0.3).

Overall, the results confirm that, in the presence of a range of different predictive models, it is possible to make efficient estimates, associated with the prediction of project delays, despite the characteristics and complexity of the data. The accuracy of the models, therefore, highlights their effectiveness and reliability in predicting delays in Italian infrastructure cohesion projects. In particular, the RF represents a very high performing algorithm in predicting project delays and, therefore, a valuable resource for policy makers in the field of territorial cohesion and projects.

In general, as already explained in several scientific works and for different sectors, for example, Kleinberg et al. (2015), Sansone (2019), and Sansone and Zhu (2023), a ML approach can have some potential as an Early Warning System (EWS). In our case, these results demonstrate how it is indeed possible to effectively predict delays in Italian infrastructure cohesion projects, thus answering our main research question affirmatively.

Table 2.5: Models' Performances

Model	RF	SVM	TREE	KNN	NN	GBM	EN	NB	Logit
Accuracy	0.829	0.795	0.779	0.773	0.767	0.767	0.744	0.700	0.700
Kappa	0.615	0.541	0.485	0.488	0.474	0.451	0.398	0.326	0.320
AccuracyLower	0.819	0.784	0.768	0.762	0.756	0.756	0.732	0.685	0.688
AccuracyUpper	0.839	0.806	0.790	0.784	0.778	0.779	0.755	0.710	0.713
AccuracyPValue	0.000	0.000	0.000	0.000	0.00	0.000	0.000	0.000	0.000
Sensitivity	0.675	0.645	0.535	0.597	0.865	0.496	0.473	0.524	0.483
Specificity	0.916	0.879	0.916	0.872	0.591	0.919	0.896	0.794	0.822
PosPredValue	0.818	0.749	0.781	0.723	0.791	0.775	0.717	0.588	0.604
NegPredValue	0.834	0.816	0.779	0.795	0.710	0.765	0.752	0.749	0.740
Precision	0.675	0.645	0.535	0.597	0.865	0.496	0.473	0.524	0.483
Recall	0.818	0.749	0.781	0.723	0.791	0.775	0.717	0.588	0.603
F1	0.740	0.693	0.635	0.654	0.826	0.605	0.570	0.554	0.537
Prevalence	0.359	0.359	0.359	0.359	0.641	0.359	0.359	0.359	0.358
DetectionRate	0.242	0.232	0.192	0.214	0.554	0.178	0.170	0.188	0.173
DetectionPrevalence	0.296	0.309	0.246	0.296	0.701	0.230	0.236	0.320	0.287
BalancedAccuracy	0.795	0.762	0.725	0.735	0.728	0.708	0.684	0.659	0.653
NoInformationRate	0.641	0.641	0.641	0.641	0.641	0.641	0.641	0.641	0.641

Note: These values are estimated on the confusion matrix. which shows a cross-tabulation of the observed and predicted classes, generating the predicted classes based on the typical 50% cut off for the probabilities (Kuhn et al., 2020). Averages were calculated across repetitions for 10 different random splits.

2.5.1.2 Features Importance

In this Section, the most important predictors of the delays of the infrastructure cohesion projects of the 2007-2013 cycle are presented (Figure 2.4). In this analysis we only use the RF which, as illustrated in the previous Section (2.5.1.1), is the algorithm that predicts delays of infrastructural cohesion projects best of all, in terms of distance of the bisector of the ROC curves and compared to the DeLong test (see Figure 2.3 and Table 2.4).

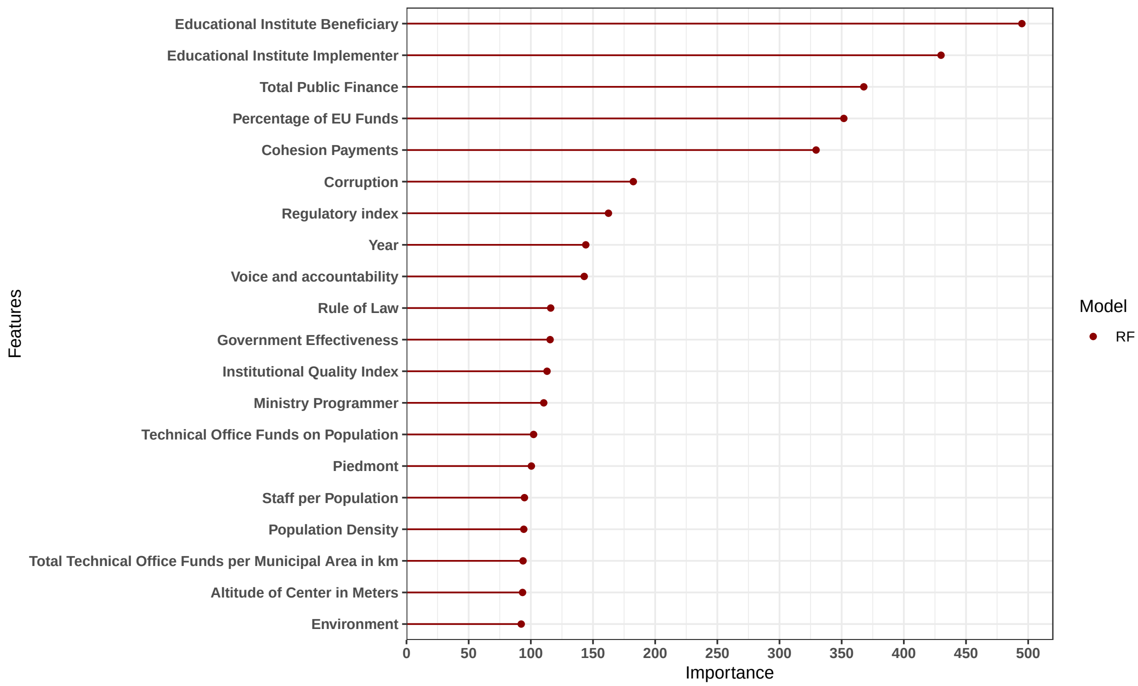
To make the most of the data available for this analysis, following established literature, e.g. Cawley and Talbot (2010), Krstajic et al. (2014), Di Stefano and Resce (2024), we trained the RF model on the entire dataset. This methodological approach ensured that every aspect of our database was exploited effectively to estimate what matters most in predicting delays.

In terms of the relative importance of features (see Figure 2.4), schools, both considered as implementing bodies and beneficiaries of cohesion projects, emerged as the most relevant features of the delay. This reflects the relevance of the management and internal organisation of these educational institutions, which affect the level of institution and training of the youngest (Bryk and Schneider, 2002). The efficiency with which these entities organise and manage cohesion funds and, therefore, coordinate project execution activities, can determine a significant acceleration or slowing down of the overall progress of the projects. An administrative support system, perhaps with a task force at Ministerial level, may help. The resources assigned to cohesion projects, such as the amount of public funding, the Percentage of EU Funds and the total cohesion payments obtained, are fundamental features to influence a continuous workflow of the projects. In this sense, delays in payments or mismanagement of finances can create significant bottlenecks, highlighting the need for robust and transparent financial systems to ensure that funds are available and used in a timely manner (Ferry and McMaster, 2013; Tosun, 2014; Bachtler et al., 2017). The Institutional Quality Indicator (IQI) and the composite sub-indicators of (Nifo and Vecchione, 2014), such as Corruption, Voice and Accountability, Regulatory Index and Government Effectiveness, have a significant impact on the efficiency and timing of infrastructure cohesion projects, albeit with lower intensity. For example, corruption can divert resources and slow down (D'Alpaos et al., 2013) projects. Likewise, the strength of regulatory policies and the integrity of government practices are critical to minimising delays and improving project effectiveness (Espinoza and Presbitero, 2022). Public spending on technical offices also appears to be a factor that has a significant impact on the delay (Gori et al., 2017, 2024), also in relation to the geographical and territorial characteristics of the municipalities, especially in relation to the altitude and surface area and, therefore, urbanisation (Albanese et al., 2024). In particular, the relevance of the altitude of the municipality highlights that certain territorial contexts can represent obstacles that complicate the management of projects, or that can facilitate their execution. These factors are particularly problematic in mountainous contexts, where the complexity of logistics systems can significantly influence implementation times (Davoudi and Strange, 2009). Excluding the dummy relating to projects located in the Piedmont Region, the spatial variables (regional dummies and macro regional project dummies) in this analysis of projects delays, are not among the most important features, as do the political economy variables relating to administrative elections and the characteristics of local administrators.

Overall, these feature importances confirm what has been demonstrated by the growing literature on cohesion policies, which highlights how the quality and characteristics of local institutions can undermine or improve the effectiveness and absorption of cohesion funds, for example Becker et al. (2013), Rodríguez-Pose and Garcilazo (2015), Barbero et al. (2023), Albanese et al. (2024).

In summary, the interaction of these variables creates a complex network that requires careful management of cohesion policies, to optimise project implementation times. Adopting targeted strategies to address each of these variables could not only reduce delays but also increase transparency and effectiveness in the use of cohesion funds, for completing projects on time.

Figure 2.4: Importance of Features with the Best Model (RF) in case of Classification



Note: The RF is trained on the entire database, which presents the binary variable relating to cohesion projects as an outcome variable. This training allows you to evaluate the importance of the variables in the national context. The importance of variables is determined by analysing the impact of permutation of the values of each variable on the data. This method measures the decrease in model accuracy when the values of a variable are randomly swapped, thus reflecting the importance of that variable in the prediction. Variables that cause a greater loss of accuracy when permuted are considered more important, indicating a strong impact on the predictive ability of the model. This analysis provides a clear indication of which characteristics are crucial for delays in infrastructure cohesion projects at the national level, highlighting the variables that contribute significantly to the model's predictions.

2.5.1.3 Partial Dependence

In this paragraph, we study the direction and intensity of the most important characteristics regarding project delays. To highlight the signs of the individual features, we use Partial Dependence Plots (PDP), which allow us to study the relationship between variables and delays in projects, also allowing us to detect varying dynamics and time trends[13].

The PDP relating to the characteristics of educational institutions as implementing bodies of Italian infrastructural cohesion projects (in Figure 2.5a) shows a positive

[13]For a more complete analysis of partial dependencies, following Greenwell et al. (2017), additional 2- and 3-dimensional Partial Dependence Plots are included in the Appendix (see Figures B3, B4), which explore the most relevant features.

relationship with delay. Specifically, we can observe how the probability of delay in cohesion projects increases with the involvement of educational institutions as implementing bodies. The direction and trend of this relationship is linked to the very nature of this feature (dummy variable which is equal to 1 if an educational institution implements the project or 0 if it is not). As a result, the efficiency and ability to manage complex projects worsens when educational institutions are directly involved in project implementation, contributing to increased delays. This evidence can be attributed to a poor efficiency of local institutions in managing projects not connected to their educational purposes (Accetturo et al., 2014; Incaltarau et al., 2020).

Figure 2.5b highlights, however, the relationship between total public funding for cohesion projects and their delays, revealing a relationship connected to the amount of resources assigned. Initially, in fact, increases in funding lead to a rapid increase in project delays, up to a certain resource threshold. With further increases in funding, the delay in projects remains essentially constant. This trend shows that a greater allocation of financial resources after a certain threshold does not produce significant changes in the project completion times.

When analysing the percentage of cohesion on total funding, an opposite behaviour is observed compared to total funding. In fact, Figure 2.5c highlights a positive initial trend, where an increase in the percentage of cohesion funding corresponds to a decrease in project delays. However, as in the case of total funding, beyond a certain threshold further increases in cohesion funding do not lead to significant improvements, suggesting that delays remain largely unchanged. This shows again that, although the presence of EU cohesion funds can increase the speed of execution, there is a limit beyond which increases in cohesion funding are no longer effective in further reducing delays.

The latter evidence can be determined by potential saturations in the capacity to manage funds or by other critical issues linked to the inefficiencies of cohesion projects, as highlighted in the works of Incaltarau et al. (2020), Cunico et al. (2022), Espinoza and Presbitero (2022), which suggest an overall review of the allocation of funds, to place greater focus on the optimisation of processes and management.

Other local characteristics, such as institutional quality indicators, present non-linear trends. In particular, the corruption indicator at the provincial level (see the PDP in Figure 2.5d) shows a positive relationship with delays[14]. In fact, a decrease in corruption leads to a significant reduction in delays D'Alpaos et al. (2013). Similar and more linear trends, albeit with an opposite direction in terms of the sign of the indicators (when the index is low the institutional quality is low), are also recorded in other sub-indicators of institutional quality, such as Voice and Accountability (Figure 2.4e) and Rule of Law (Figure 2.4f). In fact, they show a significant reduction in delays with the progressive increase in institutional quality,

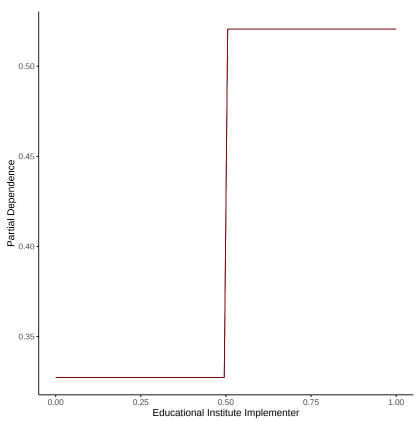
[14]High levels of corruption correspond to low values of the indicator.

up to a certain threshold, beyond which the delay begins to increase, however remaining at values lower than the initial ones. In general, these results show how several features of institutional quality are very relevant factors in determining the delay of cohesion projects, especially in the low spectrum of institutional quality, confirming the studies of Accetturo et al. (2014), and Espinoza and Presbitero (2022).

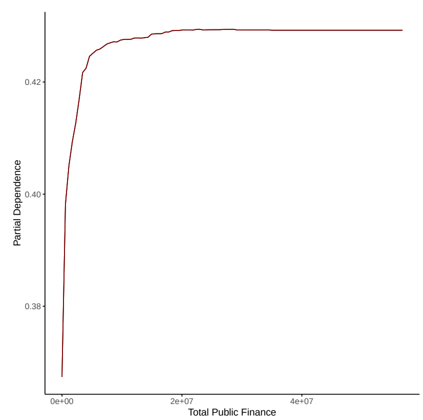
At a geographical level, the municipal altitude is a very relevant feature of the delay (see Figure 2.4g). Specifically, a positive relationship is found between delays and altitude in meters. In fact, municipal altitude determines an almost linear increase in the delay of cohesion projects, up to 100 meters (maximum delay), and then reduces slightly to remain stable in the more mountainous municipalities (Davoudi and Strange, 2009; Albanese et al., 2024).

Figure 2.5: PDP: Overview of Features Importance

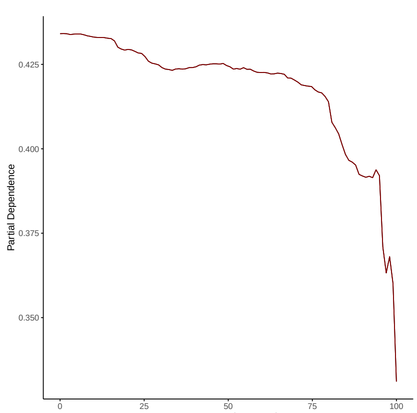
(a) Educational Institution Implementing Body



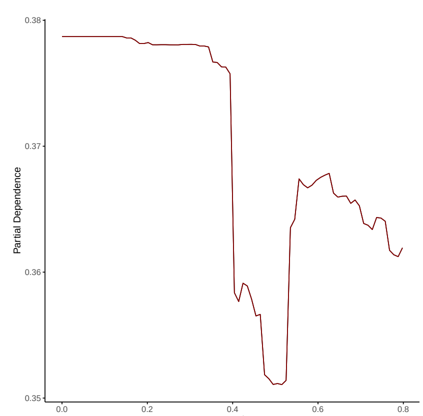
(b) Total Public Finance



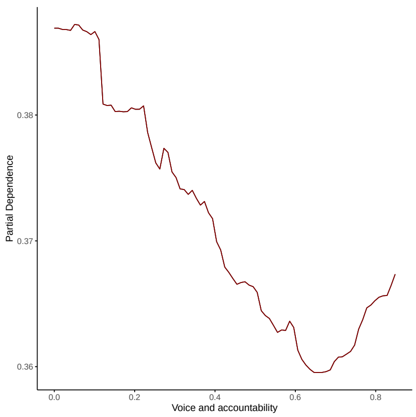
(c) Percentages of EU Funds



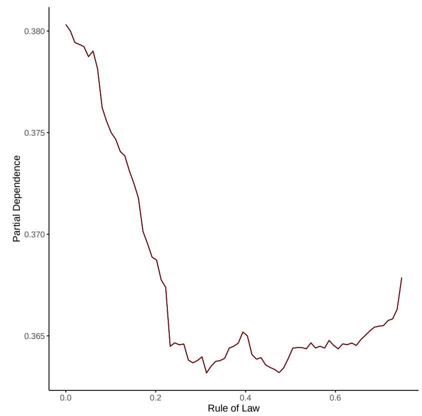
(d) Provincial Corruption Indicator



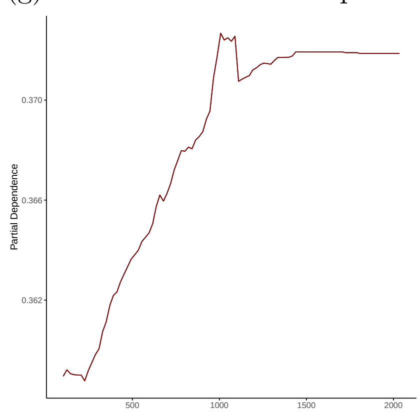
(e) Provincial Voice and Accountability Indicator



(f) Provincial Rule of Law Indicator



(g) Altitude of the Municipalities



2.5.2 Regression Task

In this Section, we present our second predictive regression task, conducted on the continuous dependent variable represented by the number of days of delay of each project.

Also, in this case, RF is the best performing algorithm for predicting delays of infrastructure projects since it has excellent performance in terms of metrics of prediction (see the values of the metrics expressed in Table 2.6.). Indeed, the RF presents the lowest values of the mean absolute error MAE[15] (on average 65.15 days of delay of the infrastructure cohesion projects), compared to the other algorithms considered, which otherwise present medium-low performances. The predictive superiority of the RF is also supported by the lowest values marked by the MSE and RSME (with an RSME of 146.54, in contrast with the values of the other algorithms between 220 and 400). Furthermore, the NN shows limited performances, indicating large errors in the prediction, making it unsuitable for regression estimates.

Table 2.6: Performance of Regression Models in Predicting Project Days of Delay

Model	RMSE	MSE	MAE
RF	146.54	21474.31	65.15
KNN	221.44	49036.99	121.26
GBM	237.1	56215.46	130.43
SVM	239.98	57590.39	123.83
TREE	250.92	62961.77	138.98
Linear model	251.34	63171.90	142.62
EN	254.39	64714.77	139.74
NN	401.34	161076.87	190.74

Note: RMSE stands for Root Square Mean Error; MSE stands for Mean squared error; MAE stands for Mean Absolute Error.

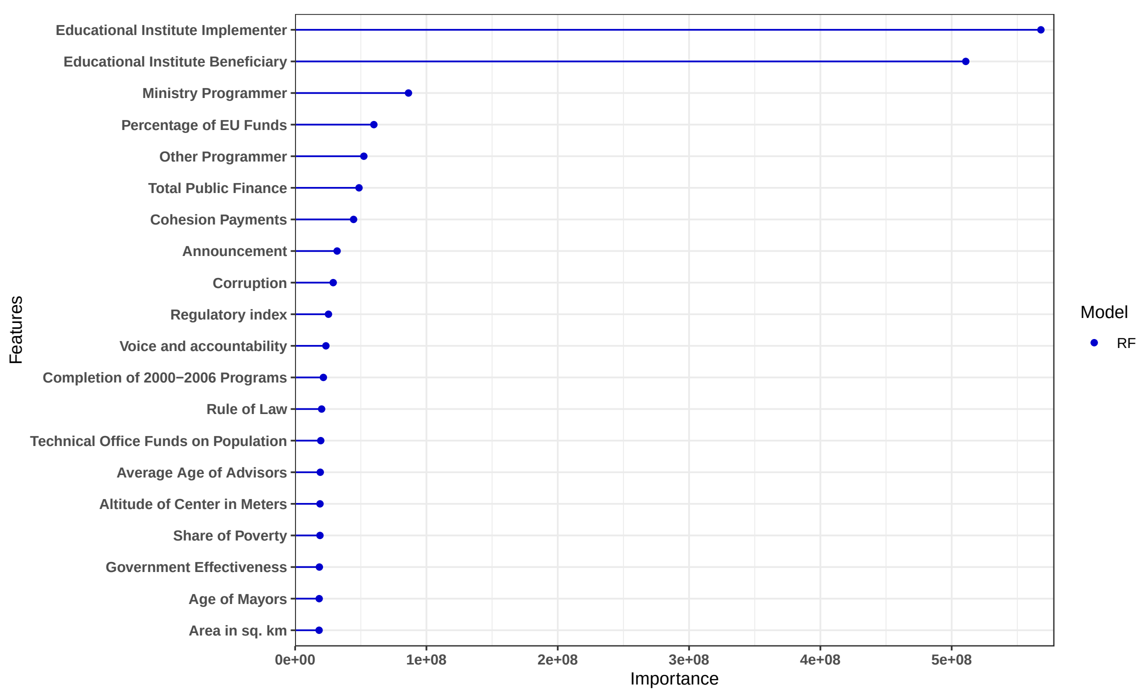
Figure 2.5 reports the features' importance in the case of regression. The importance of the features with a continuous variable follows that recorded in the classification analysis with a binary variable (see Figure 2.4), with a correlation between the feature importance of these two analyses that is very positive and statistically significant (of 0.734, as can be observed from the coefficients described in the note[16]). However, in this case, a significant heterogeneity among the variables is highlighted. In particular, even in this second analysis, schools appear as the implementing and beneficiary bodies that have the greatest and largest

[15] Which is a measure of accuracy used to evaluate regression models when the model predictions deviate on average from the actual values.

[16] The correlation between feature importances obtained from regression and classification models is 0.734, indicating a strong positive association between the two analyses. The 95% confidence interval for this correlation lies between 0.6610 and 0.7925, confirming the robustness of the result. The associated p-value is extremely low, 4.87×10^{-34}, suggesting that the observed correlation is statistically significant and not causal.

impact on the delays. Although of lesser importance, financial factors, such as the total public funding of each project, the percentage of funding from cohesion funds, institutional and local administration quality variables, public spending per capita for technical offices and geographical variations, are the features that most impact on project timing.

Figure 2.5: Importance of Features With the Best Model (RF) in Case of Regression



Note: The RF model is trained on the entire database, which presents as an outcome variable the continuous variable relating to the days of delay of each cohesion project considered. This training allows you to evaluate the importance of the variables in the national context.

The Partial Dependence Plots relating to the main features importance of the predictive analysis conducted using the continuous variable of the days of delay (presented in the Appendix 2.7) show very similar trends to those observed in the classification analysis (reported in Section 2.5.1.3).

In general, this second predictive analysis with ML algorithms, conducted on the variable relating to the number of days late for each project, shows very similar trends to the main analysis conducted on the late projects, confirming the validity of the estimates developed. The predictive differences obtained derive, in fact, mainly from the nature of the continuous variable, which presents a significant heterogeneity in terms of days of delay.

2.6 Territorial Heterogeneity

In this Section, we present a territorial analysis, conducted on a macro-regional scale. We do this in consideration of the marked Italian regional inequalities in socio-economic development, capable of impacting the delays of cohesion projects. Specifically, we divided our dataset between projects carried out in the municipalities of the Centre-North and the South.

The graphs in Figure 2.6 and 2.7 express the AUC of the ROC for the municipalities for the two macro-regions. Differences are found in the out-of-sample performances between the two spatial analyses, compared to the various machine algorithms learning and the identified hyperparameters.

Specifically, the performances of the trained models, considering only the infrastructure projects implemented in the Southern regions, are similar to those obtained on the total Italian dataset. Lesser performance results are obtained when considering only the projects implemented in the Centre-North.[17] However, in both these analyses, the RF model stands out as the best ML algorithm, capable of preventing delays in infrastructure cohesion projects in Italy, confirming the predictive goodness of the main estimate.

Figure 2.6: ROC Curve for the South

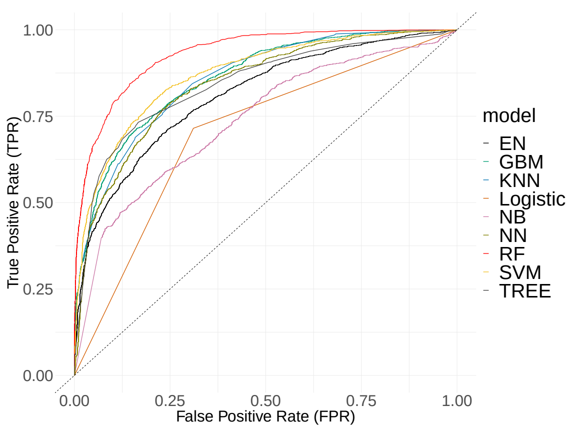
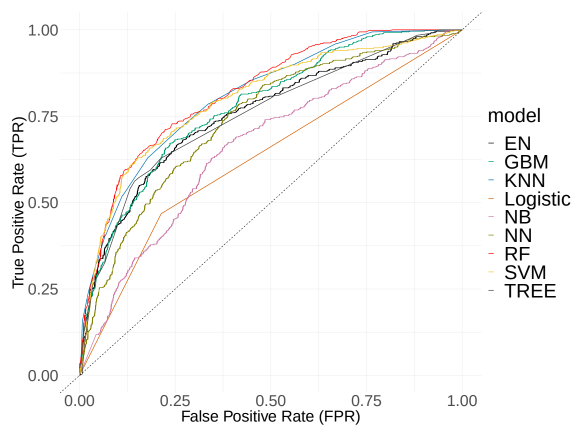


Figure 2.7: ROC Curve for the North

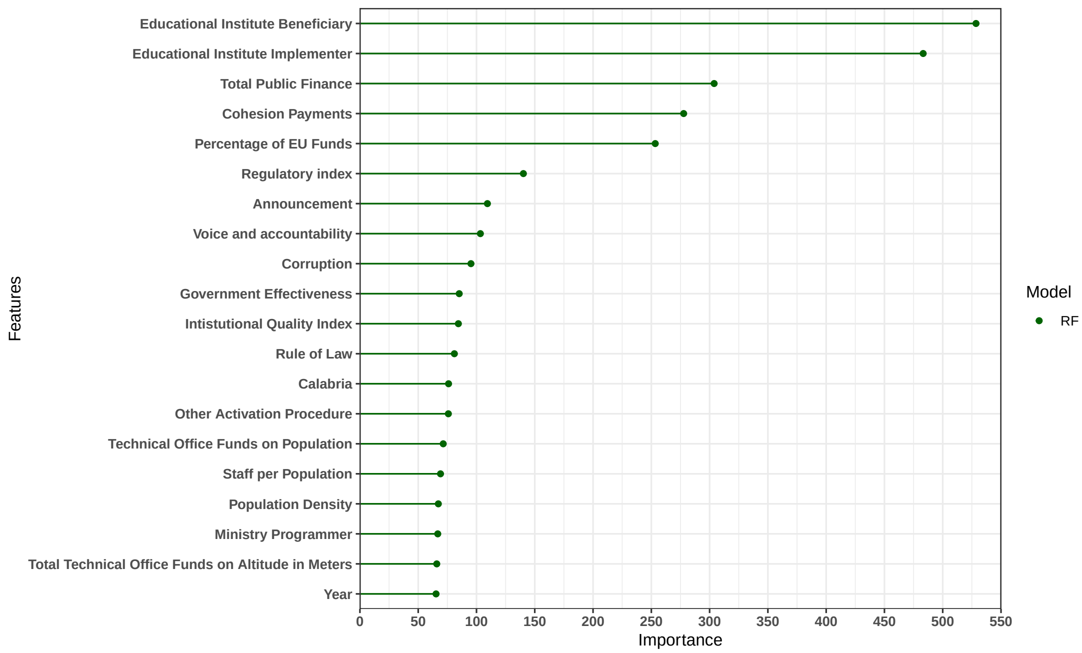


Territorial heterogeneity in delays between projects implemented in Northern and Southern Italy also emerges in the importance of features. In fact, Figures 2.8a and 2.8b highlight significant differences both in the types of features identified and in their relevance. In the analysis relating to projects in the Centre-North, the lines indicating the importance of individual features are close to each other, suggesting the significant importance of each individual feature. In contrast, in the South, a greater variation in the importance attributed to the different predictors is observed, with educational institutions emerging overwhelmingly as the main predictors.

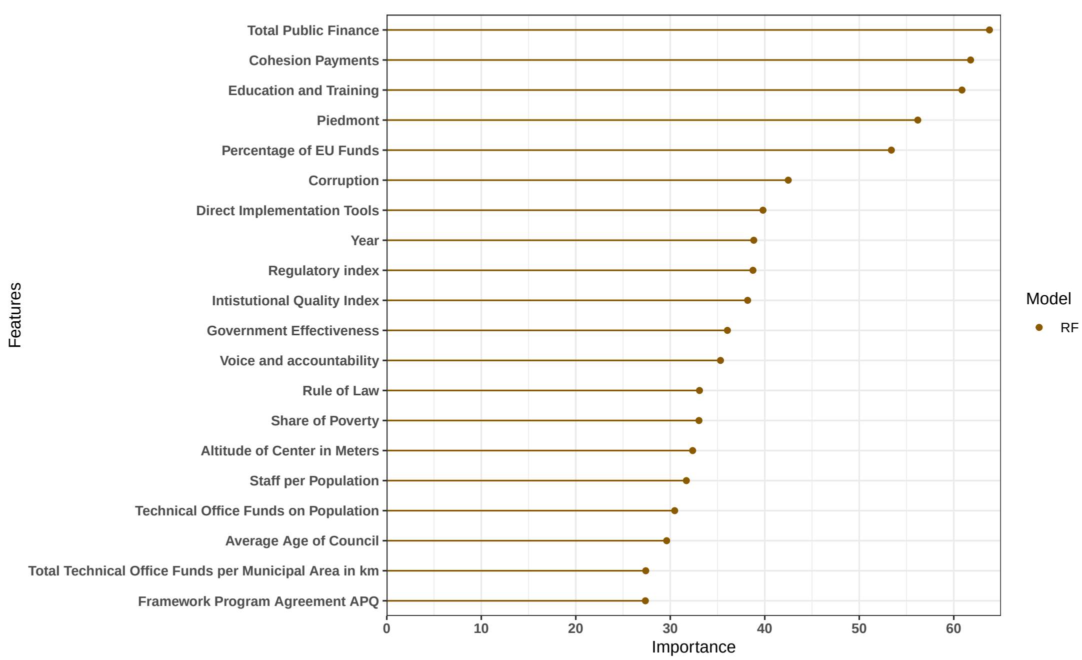
[17] Consider that there is a lower concentration of cohesion projects in these regions.

Figure 2.8: Importance of Features with the Best Model (RF) for Macro Regions

(a) South



(b) North



Note: The RF model is trained on the entire database, which presents as an outcome variable the continuous variable relating to the days of delay of each cohesion project considered. This training allows us to evaluate the importance of the variables in the South and North.

However, despite the differences that emerged in the importance of the predictors, the correlation between the feature importance for the estimates between projects implemented in the South and the North presents a very high correlation coefficient (0.797)[18], which highlights a strong relationship between the territorial analyses under examination.

In general, the features' importance in the South follows closely the results obtained in the main analysis in Section 2.5.1, with schools identified as the most important features for the completion of cohesion projects, followed by financial and management characteristics of projects and by provincial and municipal socio-institutional factors. Schools are less relevant for the conclusion of cohesion projects in the Centre-North. However, the importance of the educational sector in the timing of project implementation also emerges in projects implemented in the Centre-North, where the characteristics relating to the educational are among the most important, together with those of a financial nature.

These differences suggest important regional variations among educational systems, that may be linked to the management, and standardisation frameworks of infrastructural cohesion projects in Southern regions. Indeed, delays and the correlated inefficiency may be the main explanation for the clear disadvantage of southern schools in terms of infrastructure, as documented by Bucci et al. (2021). However, this may be linked to the quality of administrations or forms of cronyism (Coco and Lagravinese, 2014)[19], which are also connected to the general inefficiency of institutions (Neilson and Zimmerman, 2014; Coco et al., 2020).

2.7 Conclusions

This study presents an analysis of cohesion project delays, using an advanced approach based on machine learning techniques to identify and interpret the key factors for prediction. Through the use of eight different algorithms, applied to a uniquely rich database at the territorial level, we were able to isolate the most influential variables, offering a more detailed view than traditional analytical approaches.

The results demonstrate that project delays can be predicted using a set of features observed at the start date of the project. Both economic factors and institutional features play significant roles in these predictions. The most important predictor is the involvement as an implementer or beneficiary institution of a school, particularly in the less developed South. Cohesion funds are particularly concentrated in the south, making this result of tantamount importance for policy. It also explains why infrastructure endowments in this

[18] The correlation between feature importance is very high 0.797 is statistically significant (p-value: 1.79×10^{-36} with a 95% confidence interval ranging from 0.733 to 0.848).

[19] Cronyism is a non-transparent and meritless practice in which positions of power or economic opportunities are assigned and distributed to acquaintances and friends. It can damage the efficiency of institutions, leading to suboptimal decisions and inequitable distribution of resources.

sector are heterogeneous.

The use of European funds (as opposed to national co-financing) is a good predictor of the timely execution of projects. This suggests an 'advantage of tying one's hand' and offers an argument in favour of keeping cohesion policy at the EU level. It also has important consequences for the most appropriate way to implement the next budget cycle of EU funds. The importance of the EU Cohesion funds should remain intact and the discipline element, both financial and thematic, should be reinforced.

Institutional quality, in several dimensions including corruption, also emerges as a crucial determinant, suggesting that sometimes even small improvements in these areas could facilitate greater time-efficiency of cohesion programmes. Our work, in general, highlights the importance of considering a variety of factors, including less tangible ones such as the quality of local governance. Our models revealed that some local dynamics, administrative capacities and the specific demographic structure of a municipality can significantly alter project expected delays.

The implications of these findings are relevant for policymakers. Interventions aimed at improving administrative capacity and strengthening institutional quality could improve the effectiveness of the use of funds. This implies that personalised and focused support strategies may be necessary to address the specific barriers that impede the effectiveness of cohesion programmes, particularly in the education sector. Given the relevance of this sector, a centralisation of procedures on a national task force may be worth considering.

Finally, while our study has shed light on many aspects of government inaction, further research is needed to explore the complex interactions between multiple variables in different contexts and to evaluate the effectiveness of specific interventions intended to reduce delays. Key findings indicate that educational system effectiveness, administrative age profiles, and socioeconomic status are among the most significant predictors of project delays, underlining the multidimensional nature of cohesion challenges.

Moreover, EU cohesion funding seems to improve significantly the probability that a project is not late and diminish the days of delay. This seems a crucial result as we approach a new budget cycle for the EU and discuss the performance of the EU cohesion policy.

Appendix B

B1. Other Descriptive Statistics and Partial Dependence Plot

Table B1: Hyperparameters in Case of Classification

Model	IperParameters	Description	Implications
Elastic Net (EN)	Model type: $\alpha = 1$ Penalization parameter: $\lambda = 0.0003$	α at 1 indicates a pure Lasso model, favoring variable selection. α controls the degree of regularization to prevent overfitting.	A lower λ value means less regularization, which can be optimal for complex data where noise reduction is not a primary concern.
Random Forest (RF)	Number of variables per split: $mtry = 98$	$mtry$ determines how many variables are examined at each split. A high value allows for greater diversity in the tree models.	A high $mtry$ can improve accuracy but increase the risk of overfitting if the dataset is not sufficiently large or varied.
Gradient Boosting Machine (GBM)	Total number of trees: $n.trees = 150$ Interaction depth: $interaction.depth = 3$ Learning rate: $shrinkage = 0.1$ Minimum observations per node: $n.minobsinnode = 10$	$n.trees$ is the number of sequential trees. Interaction.depth manages the depth of the tree. $Shrinkage$ reduces the contribution of each tree. $n.minobsinnode$ sets the minimum observations per node.	$n.trees = 150$ with a shallow depth and shrinkage ensure gradual learning while avoiding overfitting the training data.
Neural Network (NN)	Number of units in hidden layer: $Size = 5$ Weight decay: $Decay = 0.1$	$Size$ indicates the number of neurons in the hidden layer which increases the model's ability to learn complex relationships. $Decay$ helps to reduce overfitting by penalizing larger weights.	5 neurons offer good complexity without being too computationally demanding, while a decay of 0.1 prevents the model from overfitting the training data.
K-Nearest Neighbors (KNN)	Number of neighbors considered: $k = 15$	k represents the number of nearest neighbors considered for classification or regression.	A larger k provides more stabilized decisions that are less sensitive to fluctuations in the input data.
Support Vector Machine (SVM)	Gaussian kernel width: $\sigma = 0.0043$; Error penalization parameter: $c = 8$	σ influences the width of the Gaussian RBF kernel, c is the error penalization parameter.	A smaller leads to a finer decision boundary, while a high c severely punishes misclassifications, emphasizing correct classification.
Decision Tree (TREE)	Tree pruning complexity parameter: $cp = 0.0026$	cp is the complexity parameter that regulates the growth of the tree.	A lower cp allows more complex trees, potentially capturing better the relationships in the data, but at the risk of overfitting.
Naive Bayes (NB)	Laplace correction index: $Laplace = 0$ Use of kernel approach for density: $UseKernel$ = FALSE Smoothing parameter: Adjust = 1	$Laplace$ and $UseKernel$ affect probability handling, $Adjust$ modifies the kernel parameters for continuous estimates.	The configuration avoids the use of kernel, preferring a simpler and more direct approach to probabilities, useful in situations with a good data distribution.

Table B2: Descriptive Statistics for the Features

Outcome variables	**Mean**	**SD**	**Median**	**Min**	**Max**	**N.Obs.**	**Variable**	**Surch**
- For Classification								
Dummy Delay	0.38	0.49	0	0	1	22162	Dummy	Opencoesione.it
- For Regression								
Days Delay	171.86	395.02	0	-2372	2754	22162	Continues	Opencoesione.it
Endugenus on the Opencoesione dataset								
Business Competitiveness	0.01	0.09	0	0	1	22162	Dummy	Opencoesione.it
Circular	0.01	0.09	0	0	1	22162	Dummy	Opencoesione.it
Administrative capacity	0	0.04	0	0	1	22162	Dummy	Opencoesione.it
Announcement	0.26	0.44	0	0	1	22162	Dummy	Opencoesione.it
Cohesion Payments	327754.61	1098403.28	99548.86	0	51312827.85	22162	Continues	Opencoesione.it
Company Beneficiary	0.02	0.15	0	0	1	22162	Dummy	Opencoesione.it
Company Implementer	0.02	0.15	0	0	1	22162	Dummy	Opencoesione.it
Completion of 2000-2006 Programs	0.01	0.1	0	0	1	22162	Dummy	Opencoesione.it
Culture and Tourism	0.24	0.43	0	0	1	22162	Dummy	Opencoesione.it
Digital Networks and Services	0.02	0.12	0	0	1	22162	Dummy	Opencoesione.it
Direct Identification in the Program	0.08	0.28	0	0	1	22162	Dummy	Opencoesione.it
Direct Implementation Tools	0.26	0.44	0	0	1	22162	Dummy	Opencoesione.it
Education and Training	0.18	0.39	0	0	1	22162	Dummy	Opencoesione.it
Education Training and Supports for the Labor Market	0	0.03	0	0	1	22162	Dummy	Opencoesione.it
Educational Institute Beneficiary	0.13	0.33	0	0	1	22162	Dummy	Opencoesione.it
Educational Institute Implementer	0.13	0.33	0	0	1	22162	Dummy	Opencoesione.it
EM	0.01	0.11	0	0	1	22162	Dummy	Opencoesione.it
Employment and Labor	0	0.06	0	0	1	22162	Dummy	Opencoesione.it
Energy	0.17	0.37	0	0	1	22162	Dummy	Opencoesione.it
Energy Sector Infrastructure	0.07	0.25	0	0	1	22162	Dummy	Opencoesione.it
Enrolled from Abroad	322.91	1575.02	27	0	37262	22162	Dummy	Opencoesione.it
Enrolled from Inside	793.13	2797.65	110	0	65213	22162	Dummy	Opencoesione.it
Environment	0.26	0.44	0	0	1	22162	Dummy	Opencoesione.it
Environmental Infrastructure and Water Resources	0.36	0.48	0	0	1	22162	Dummy	Opencoesione.it
ERDF (European Regional Development Fund)	0.59	0.49	1	0	1	22162	Dummy	Opencoesione.it
ESF (European Social Fund)	0	0.02	0	0	1	22162	Dummy	Opencoesione.it
Extraordinary Maintenance Restoration and Renovation	0.63	0.48	1	0	1	22162	Dummy	Opencoesione.it
Framework Program Agreement APQ	0.14	0.35	0	0	1	22162	Dummy	Opencoesione.it
Infrastructure for Telecommunications and Information Technologies	0.01	0.11	0	0	1	22162	Dummy	Opencoesione.it
Infrastructure for the Equipment of Productive Areas	0.01	0.1	0	0	1	22162	Dummy	Opencoesione.it
Ministry Beneficiary	0.02	0.13	0	0	1	22162	Dummy	Opencoesione.it
Ministry Implementer	0.02	0.13	0	0	1	22162	Dummy	Opencoesione.it
Ministry Programmer	0.21	0.41	0	0	1	22162	Dummy	Opencoesione.it
Mountain Community Beneficiary	0.17	0.38	0	0	1	22162	Dummy	Opencoesione.it
Mountain Community Implementer	0.17	0.38	0	0	1	22162	Dummy	Opencoesione.it
Municipal Beneficiary	0.39	0.49	0	0	1	22162	Dummy	Opencoesione.it
Municipal Implementer	0.4	0.49	0	0	1	22162	Dummy	Opencoesione.it
Municipal Programmer	0.01	0.09	0	0	1	22162	Dummy	Opencoesione.it
Negotiation Procedure	0.22	0.42	0	0	1	22162	Dummy	Opencoesione.it
New Realization	0.17	0.37	0	0	1	22162	Dummy	Opencoesione.it
Ordinary Maintenance	0.07	0.25	0	0	1	22162	Dummy	Opencoesione.it
Other Activation Procedure	0.27	0.44	0	0	1	22162	Dummy	Opencoesione.it
Other Beneficiary	0.45	0.5	0	0	1	22162	Dummy	Opencoesione.it
Other different from ERDF ESF	0.41	0.49	0	0	1	22162	Dummy	Opencoesione.it
Other Implementer	0.44	0.5	0	0	1	22162	Dummy	Opencoesione.it
Other Intervention	0.14	0.34	0	0	1	22162	Dummy	Opencoesione.it
Other Programmer	0.22	0.41	0	0	1	22162	Dummy	Opencoesione.it
Percentage EU Funds	88.65	20.7	99.01	0	100	22162	Continues	Opencoesione.it
Private Sector Beneficiary	0.05	0.21	0	0	1	22162	Dummy	Opencoesione.it
Private Sector Implementer	0.05	0.21	0	0	1	22162	Dummy	Opencoesione.it
Private Sector Programmer	0	0.05	0	0	1	22162	Dummy	Opencoesione.it
Projects originated from other implementing tools QSN	0	0.07	0	0	1	22162	Dummy	Opencoesione.it
Provincial Beneficiary	0.09	0.28	0	0	1	22162	Dummy	Opencoesione.it
Provincial Implementer	0.09	0.28	0	0	1	22162	Dummy	Opencoesione.it
Provincial Programmer	0.02	0.15	0	0	1	22162	Dummy	Opencoesione.it
Public Entity Beneficiary	0.02	0.15	0	0	1	22162	Dummy	Opencoesione.it
Public Entity Implementer	0.02	0.15	0	0	1	22162	Dummy	Opencoesione.it
Public Notice	0.09	0.29	0	0	1	22162	Dummy	Opencoesione.it
Public Sector Beneficiary	0.94	0.23	1	0	1	22162	Dummy	Opencoesione.it

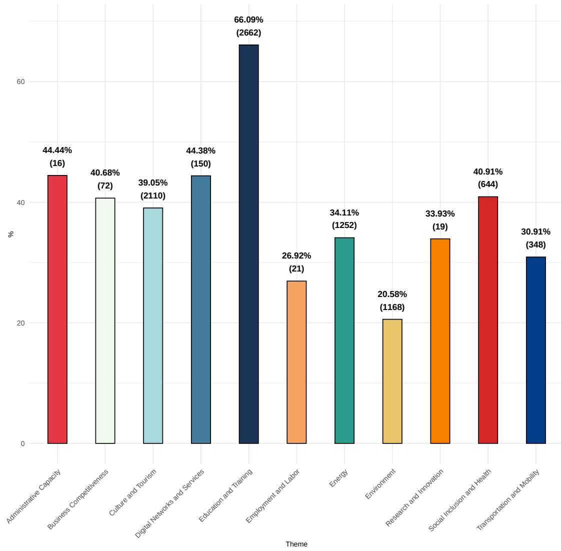
Public Sector Implementer	0.95	0.22	1	0	1	22162	Dummy	Opencoesione.it
Public Sector Programmer	1	0.05	1	0	1	22162	Dummy	Opencoesione.it
Public-Private Sector Beneficiary	0	0.03	0	0	1	22162	Dummy	Opencoesione.it
Public-Private Sector Implementer	0	0.03	0	0	1	22162	Dummy	Opencoesione.it
Regional Beneficiary	0.08	0.27	0	0	1	22162	Dummy	Opencoesione.it
Regional Implementer	0.08	0.27	0	0	1	22162	Dummy	Opencoesione.it
Regional Programmer	0.75	0.43	1	0	1	22162	Dummy	Opencoesione.it
Research and Innovation	0	0.05	0	0	1	22162	Dummy	Opencoesione.it
Research Technological Development and Innovation	0	0.04	0	0	1	22162	Dummy	Opencoesione.it
Retroactive Identification Procedure	0.05	0.21	0	0	1	22162	Dummy	Opencoesione.it
Social Inclusion and Health	0.07	0.26	0	0	1	22162	Dummy	Opencoesione.it
Social Infrastructure	0.46	0.5	0	0	1	22162	Dummy	Opencoesione.it
Total Beneficiaries	1.02	0.35	1	1	28	21997	Categorical	Opencoesione.it
Total Implementers	1.02	0.35	1	1	28	22148	Categorical	Opencoesione.it
Total Programmers	1	0.07	1	1	4	22162	Categorical	Opencoesione.it
Total Public Finance	394078.27	1621302.39	123286.04	171.45	156797676	22162	Dummy	Opencoesione.it
Total Rounds	1.14	0.35	1	1	2	19388	Dummy	Opencoesione.it
Transport Infrastructure	0.09	0.29	0	0	1	22162	Dummy	Opencoesione.it
Transportation and Mobility	0.05	0.22	0	0	1	22162	Dummy	Opencoesione.it
Variation of Structural Funds related to 2007-2013 programming	0.58	0.49	1	0	1	22162	Dummy	Opencoesione.it
Variation of the Cohesion Action Plan - Own Resources	0.13	0.33	0	0	1	22162	Dummy	Opencoesione.it
Variation of the Fund for Development and Cohesion	0.3	0.46	0	0	1	22162	Dummy	Opencoesione.it
Works, Systems And Equipment for Production Activities, Research and Social Enterprise	0.01	0.09	0	0	1	22162	Dummy	Opencoesione.it
Year	2011	2.07	2011	2007	2015	22162	Categorical	Opencoesione.it
Socio-Economic								
Per capita Income	14622.92	3453.19	13909.9	7103.54	51403.21	22162	Continues	Ministry of the Economy
Share of Poverty	0.42	0.11	0.43	0.03	0.76	22157	Continues	Ministry of the Economy
Demographic								
Births	394.66	1414.98	43	0	25958	22162	Continues	ISTAT
Cancellations for Abroad	108.79	507.36	15	0	15301	22162	Continues	ISTAT
Cancellations for Domestic	915.42	3218.09	124	0	47195	22162	Continues	ISTAT
Deaths	449.43	1591.52	54	0	27590	22162	Continues	ISTAT
Population at the beginning of the period	43566.62	151175.09	5186	47	2724565	22162	Continues	ISTAT
End-of-Period Population	43603.68	151537.2	5172	45	2760303	22162	Continues	ISTAT
Population between 50 and 100 k	0.07	0.26	0	0	1	22162	Continues	ISTAT
Population over 100 k	0.07	0.26	0	0	1	22162	Continues	ISTAT
Population Density	529.17	1215.16	133.68	0.86	12234.78	22113	Continues	ISTAT
Territorial dummies								
Center of Italy	0.1	0.3	0	0	1	22162	Dummy	ISTAT
Northern Italy	0.17	0.37	0	0	1	22162	Dummy	ISTAT
Southern Italy	0.73	0.44	1	0	1	22162	Dummy	ISTAT
Abruzzo	0.03	0.18	0	0	1	22162	Dummy	ISTAT
Basilicata	0.05	0.22	0	0	1	22162	Dummy	ISTAT
Calabria	0.09	0.29	0	0	1	22162	Dummy	ISTAT
Campania	0.27	0.44	0	0	1	22162	Dummy	ISTAT
Lazio	0.01	0.12	0	0	1	22162	Dummy	ISTAT
Liguria	0.04	0.2	0	0	1	22162	Dummy	ISTAT
Lombardy	0.03	0.16	0	0	1	22162	Dummy	ISTAT
Marche	0.03	0.16	0	0	1	22162	Dummy	ISTAT
Molise	0.02	0.13	0	0	1	22162	Dummy	ISTAT
Piedmont	0.04	0.2	0	0	1	22162	Dummy	ISTAT
Apulian	0.09	0.29	0	0	1	22162	Dummy	ISTAT
Sardinia	0.03	0.18	0	0	1	22162	Dummy	ISTAT
Sicily	0.15	0.35	0	0	1	22162	Dummy	ISTAT
Umbria	0.03	0.16	0	0	1	22162	Dummy	ISTAT
Veneto	0.02	0.14	0	0	1	22162	Dummy	ISTAT
Tuscany	0.03	0.17	0	0	1	22162	Dummy	ISTAT
Geographical								
Altimetric Zone	3.05	1.44	3	1	5	22162	Categorical	ISTAT
Altitude of Center in Meters	342.98	276.66	303	0	2035	22162	Continues	ISTAT
Area in sq. km	75.29	94.5	40.43	0.12	1287.39	22113	Continues	ISTAT
Coastal Municipality	0.2	0.4	0	0	1	22162	Dummy	ISTAT
Coastal Zones	1.52	1.19	1	0	3	22162	Dummy	ISTAT
Island Municipality	0.13	0.33	0	0	1	22162	Dummy	ISTAT
Institutional Quality and Government Effectiveness								
Intistutional Quality Index	0.4	0.24	0.38	0	1	22162	Dummy	Opencoesione.it
Corruption	0.7	0.22	0.71	0	1	22162	Dummy	Ministry of Interior
Government Effectiveness	0.32	0.15	0.31	0	1	22162	Continues	Ministry of Interior
Voice and accountability	0.43	0.22	0.4	0	1	22162	Continues	Ministry of Interior
Rule of Law	0.4	0.22	0.35	0	1	22162	Continues	Ministry of Interior
Regulatory index	0.41	0.22	0.39	0	1	22162	Continues	Ministry of Interior

Local Policy and Administrative								
Age of Mayors	65.13	9.66	65	32	108	22153	Dummy	Ministry of Interior
Average Age of Advisors	58.82	4.61	59	39.8	76.42	22162	Dummy	Ministry of Interior
Average Age of Council	60.84	5.81	61	37	90.67	22162	Dummy	Ministry of Interior
Centre	0.01	0.09	0	0	1	19359	Dummy	Ministry of Interior
Centre-left	0.16	0.36	0	0	1	19359	Dummy	Ministry of Interior
Centre-right	0.16	0.36	0	0	1	19359	Dummy	Ministry of Interior
Civic List Text Analysis	0.19	0.39	0	0	1	22162	Dummy	Ministry of Interior
Civic Lists	0.68	0.47	1	0	1	19359	Dummy	Ministry of Interior
Coalition List Votes in I Round out of Voters	52.61	25.81	52.11	11.09	2952.79	19359	Dummy	Ministry of Interior
Coalition List Votes in I Rouout out of Electors	37.85	16.53	37.16	6.76	1698.73	19359	Dummy	Ministry of Interior
Election Year	2008.95	2.9	2009	2003	2015	19388	Categorical	Ministry of Interior
Five Stars MoViment	0	0.03	0	0	1	19388	Dummy	Ministry of Interior
Lists Over 60 Electors	0.03	0.16	0	0	1	19359	Dummy	Ministry of Interior
Lists Over 60 Voters	0.23	0.42	0	0	1	19359	Dummy	Ministry of Interior
Lists Over 80 Electors	0	0.04	0	0	1	19359	Dummy	Ministry of Interior
Lists Over 80 Voters	0.04	0.2	0	0	1	19359	Dummy	Ministry of Interior
Mayor Gender	0.07	0.26	0	0	1	22162	Dummy	Ministry of Interior
Mayor Over 60 Electors	0.02	0.14	0	0	1	19359	Dummy	Ministry of Interior
Mayor Over 60 Voters	0.22	0.42	0	0	1	19359	Dummy	Ministry of Interior
Mayor Over 80 Voters	0.04	0.19	0	0	1	19359	Dummy	Ministry of Interior
Mayors' Education Level	0.62	0.48	1	0	1	22162	Dummy	Ministry of Interior
Mayors Over 50	0.93	0.26	1	0	1	22153	Dummy	Ministry of Interior
Mayors Over 70	0.31	0.46	0	0	1	22153	Dummy	Ministry of Interior
Mayors Under 40	0	0.05	0	0	1	22153	Dummy	Ministry of Interior
Number Assessors	5.33	2.77	4	0	17	22039	Continues	Ministry of Interior
Number Councilors	18.45	10.5	16	6	60	22039	Continues	Ministry of Interior
Reelected Mayor in I Term after 2007	0.8	0.4	1	0	1	19388	Dummy	Ministry of Interior
Reelected Mayor in II Term after 2007	0.2	0.4	0	0	1	19388	Dummy	Ministry of Interior
Victorious in I Round Without Runoff	0.13	0.33	0	0	1	19388	Dummy	Ministry of Interior
Victorious in I Round	0.86	0.35	1	0	1	19388	Dummy	Ministry of Interior
Victorious in II Round	0.14	0.35	0	0	1	19388	Dummy	Ministry of Interior
Voters	34142.4	127152.1	4576	70	2359119	19359	Continues	Ministry of Interior
Votes	23878.32	85474.54	3340	49	1729287	19359	Continues	Ministry of Interior
Votes for Head Candidate in I Round Electors	37.98	9.77	37.53	10.09	86.17	19359	Dummy	Ministry of Interior
Votes for Head Candidate in I Round out of Voters	52.75	12.57	52.07	16.48	98.32	19359	Dummy	Ministry of Interior
Votes for Head Candidate in I/II Round out of Electors	38.6	9.17	38	12.95	86.17	19359	Dummy	Ministry of Interior
Votes for Head Candidate in I/II Round out of Voters	53.6	11.61	52.2	19.2	98.32	19359	Dummy	Ministry of Interior
Technical Office Funds								
Purchase of Goods and/or Raw Materials	16639.51	68808.69	2987.5	0	2832897	22162	Continues	Openbilanci.it
Purchase of Goods and/or Raw Materials on Percentage EU Funds	0.36	6.26	0.01	0	566.16	21728	Continues	Openbilanci.it
Purchase of Goods and/or Raw Materials on Total Financing	0.3	5.31	0.01	0	427.45	22162	Continues	Openbilanci.it
Purchase of Goods and/or Raw Materials per Altitude in Meters	434.84	6938.8	11.87	0	809620.8	22150	Continues	Openbilanci.it
Purchase of Goods and/or Raw Materials per Municipal Area in km	699.42	5374.97	62.1	0	490100.73	22116	Continues	Openbilanci.it
Purchase of Goods and/or Raw Materials per Per Capita Income	1.2	5.18	0.21	0	240.35	22162	Continues	Openbilanci.it
Purchase of Goods and/or Raw Materials per Population	8.06	75.98	0.38	0	6283.99	22162	Continues	Openbilanci.it
Purchase of Goods and/or Raw Materials per Share of Poverty	42208.51	175750.22	7175.08	0	5934025.11	22162	Continues	Openbilanci.it
Staff	617247.1	1807829.6	150866	0	35814598.2	22162	Continues	Openbilanci.it
Staff per Altitude in Meters	16052.43	138656.3	687.16	0	10091560	22154	Continues	Openbilanci.it
Staff per Population	280.59	1809.98	29.02	0	111535.25	22162	Continues	Openbilanci.it
Staff per Per Capita Income	44.43	134.47	10.75	0	2938.59	22162	Continues	Openbilanci.it
Staff per Share of Poverty	1568028.1	4708987.8	375110.1	0	109870066	22162	Continues	Openbilanci.it
Staff per Municipal Area in km	24164.17	107277.48	3884.38	0	4853272.98	22110	Continues	Openbilanci.it
Staff on Total Financing	14.36	458.29	0.7	0	61659.81	22162	Continues	Openbilanci.it
Staff on Total Cohesion Payments	18.3	597.71	0.79	0	78128.55	21724	Continues	Openbilanci.it
Service Performance	171640.2	1825730.9	15036	0	84394852.6	22162	Continues	Openbilanci.it
Service Performance per Altitude in Meters	5006.47	168139.3	63.42	0	23266960	22154	Continues	Openbilanci.it
Service Performance per Population	74.46	1155.92	2.19	0	93262.77	22162	Continues	Openbilanci.it
Service Performance per Per Capita Income	12.21	133.98	1.05	0	7222.25	22162	Continues	Openbilanci.it
Service Performance per Share of Poverty	440105.58	4772727.56	36345.39	0	270645995.47	22162	Continues	Openbilanci.it
Service Performance per Municipal Area in km	6872.76	118792.8	330.97	0	15043650	22110	Continues	Openbilanci.it
Service Performance on Total Financing	4.24	250.13	0.06	0	36588.74	22162	Continues	Openbilanci.it
Service Performance on Total Cohesion Payments	5.24	263.95	0.07	0	36977.66	21724	Continues	Openbilanci.it
Technical Office Funds on Population	426.03	3567.7	37.29	0	288676.48	22162	Continues	Openbilanci.it
Total Technical Office Funds	919352.1	3864733.5	190036	0	160484955.6	22162	Continues	Openbilanci.it
Total Technical Office Funds on Altitude in Meters	21181.37	167975.68	912.66	0	8321063.1	22150	Continues	Openbilanci.it
Total Technical Office Funds on Total Financing	15.18	279.04	0.91	0	37003.87	22162	Continues	Openbilanci.it
Total Technical Office Funds per Municipal Area in km	38153.5	276035.75	5070.61	0	29885466.59	22116	Continues	Openbilanci.it
Total Technical Office Funds per Per Capita Income	66.31	283.08	13.53	0	12589.45	22162	Continues	Openbilanci.it
Total Technical Office Funds per Share of Poverty	2336397.7	10183916.9	481043.6	0	495347415.3	22162	Continues	Openbilanci.it
Total Technical Office Funds on Cohesion Payments	17.71	306.89	1.02	0	39430.16	21728	Continues	Openbilanci.it

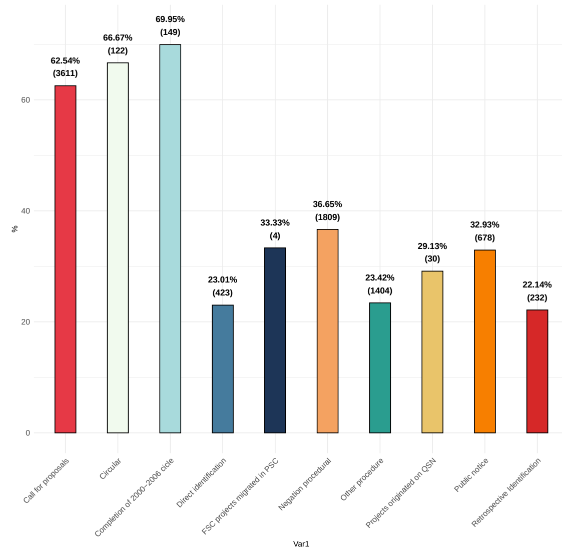
Source: Authors' elaborations.

Figure B1: Infrastructure Cohesion Projects Delayed Across Various Administrative and Procedural Aspects

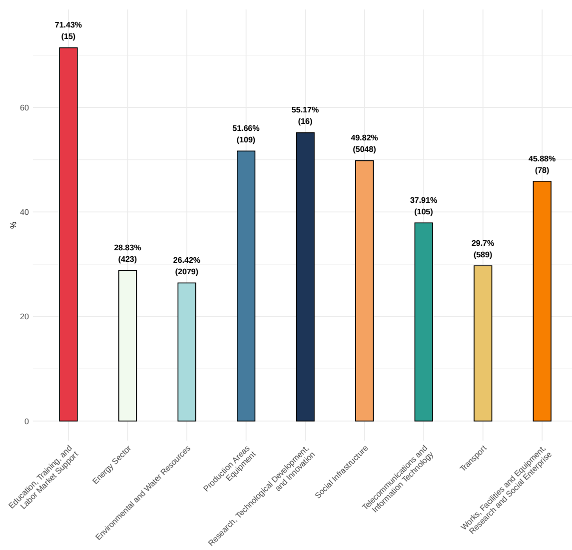
(a) By Themes



(b) By Procedures and Execution



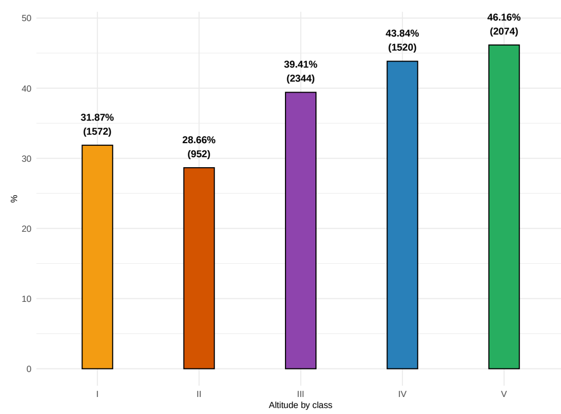
(c) By Economic Sectors



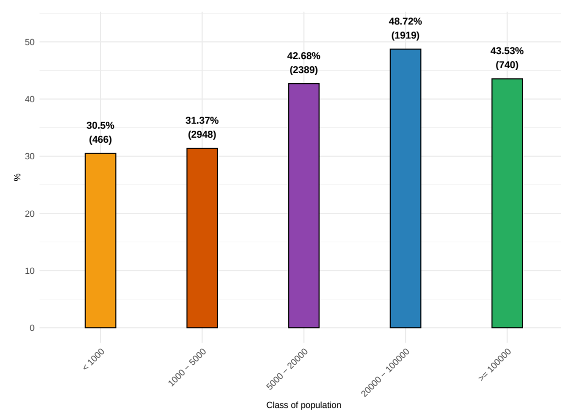
Source: Authors' elaborations on Opencoesione.

Figure B2: Infrastructure Projects Delayed at Municipal Level

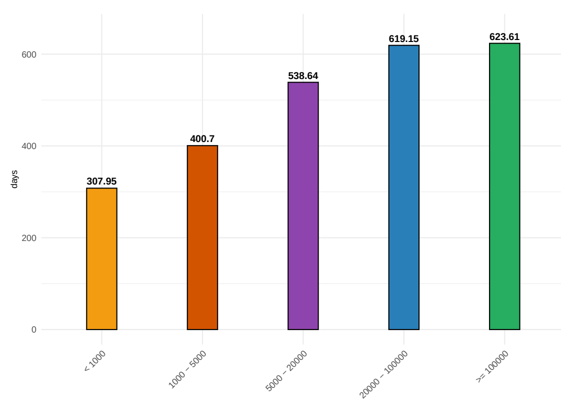
(a) Delays in Projects Relative to the Altitude of Municipalities



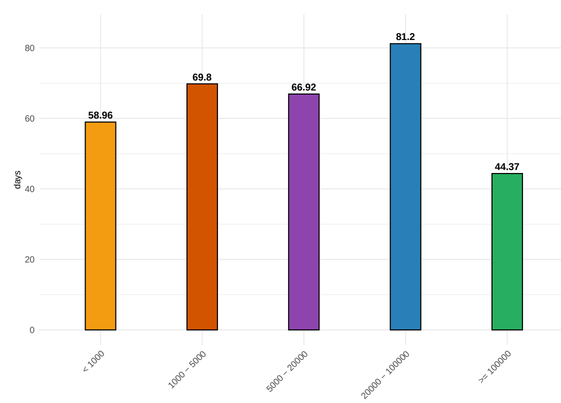
(b) Project Daleyd by Municipal Population Classes



(c) Average Days of Delay by Municipal Population Classes



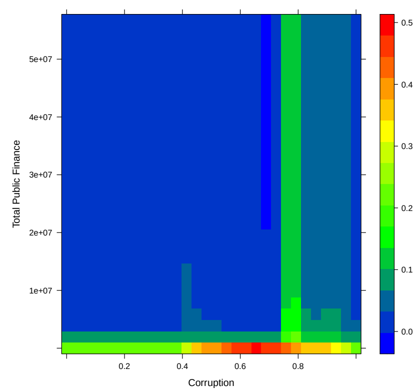
(d) Average Days of Delay by Municipal Population Classes when Municipalities are Implementing Body



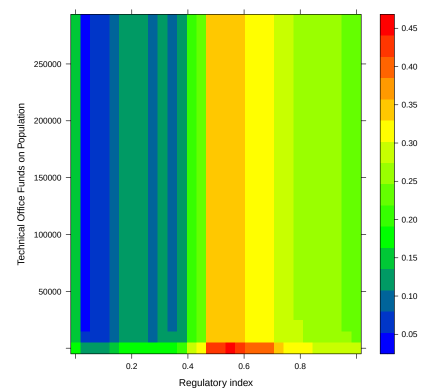
Source: Authors' elaborations on Opencoesione.

Figure B3: 2-Dimension PdP

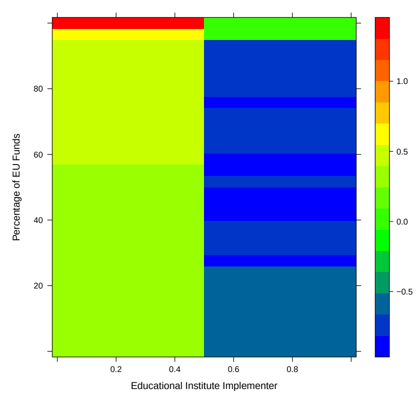
(a) Corruption and Total Public Finance



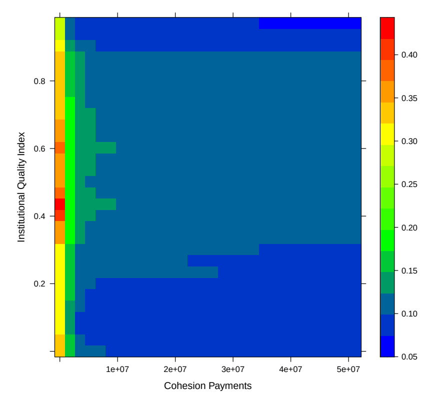
(b) Regulatory Index and TOP/Pop



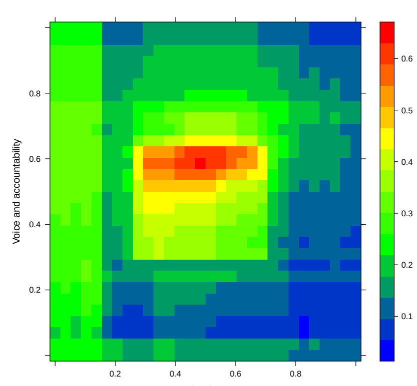
(c) Educational Institute and EU-Funds



(d) Cohesion Payments and IQI



(e) Rule of Law and Voice-Accountability



(f) Altitude in M. and Pop. Density

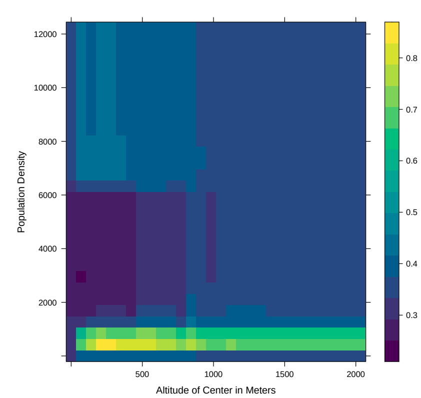
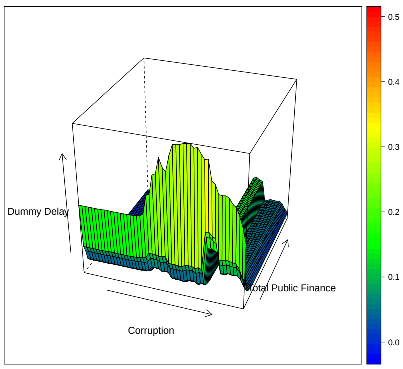
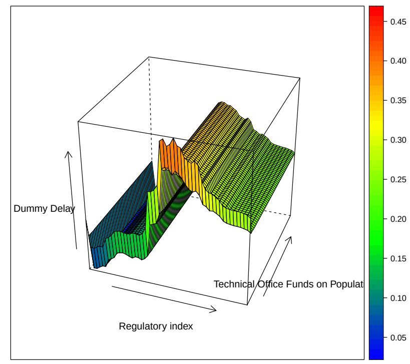


Figure B4: 3-Dimension PdP

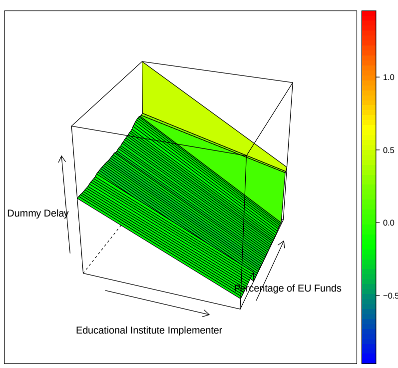
(a) Corruption and Total Public Finance



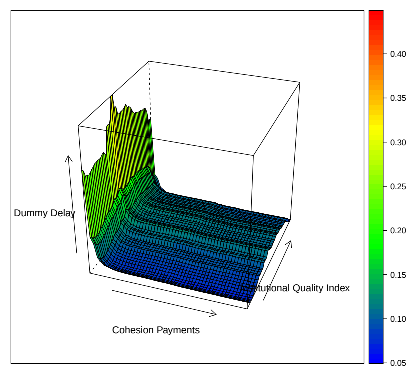
(b) Regulatory Index and TOP/Pop



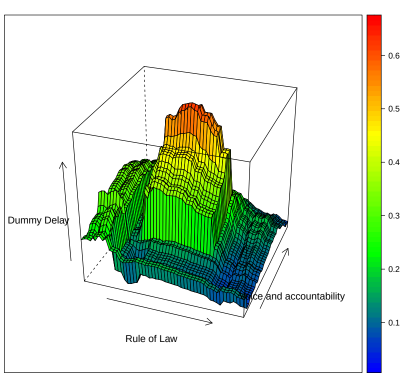
(c) Educational Institute and EU-Funds



(d) Cohesion Payments and IQI



(e) Rule of Law and Voice-Accountability



(f) Altitude in m. and Pop. Density

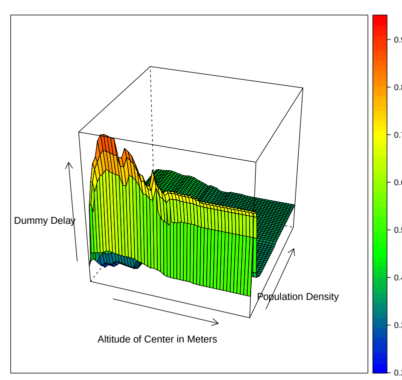
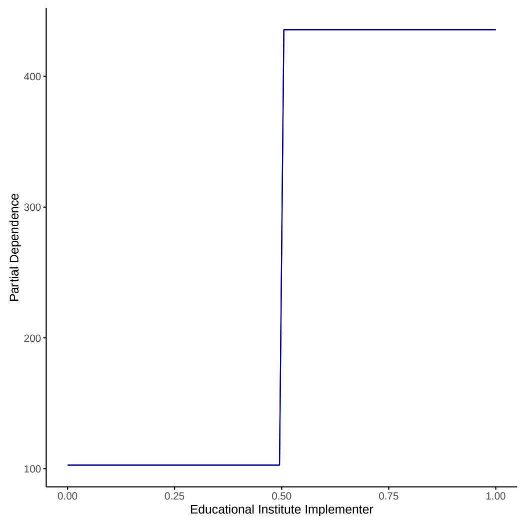
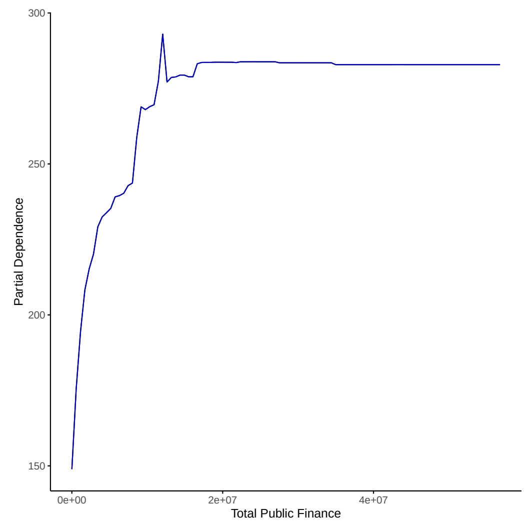


Figure B5: PDP: Overview of Features Importance in Case of Regression

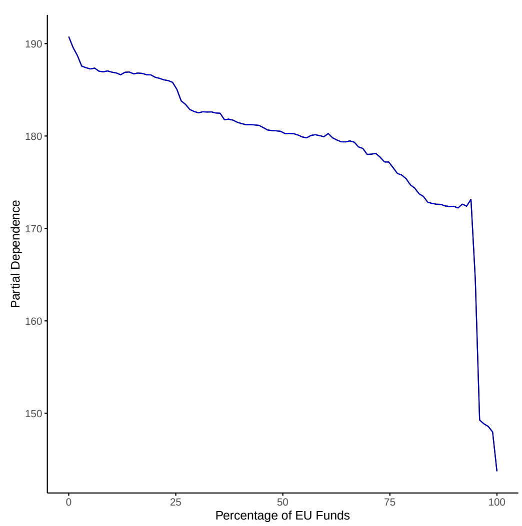
(a) PDP: Educational Institution Implementing Body in case of Regression



(b) PDP: Total Public Finance in case of Regression



(c) PDP: Percentages of EU Funds in case of Regression



Bibliography

Accetturo, A. and De Blasio, G. (2012). Policies for local development: An evaluation of italy's "patti territoriali". *Regional Science and Urban Economics*, 42(1-2):15–26.

Accetturo, A., De Blasio, G., and Ricci, L. (2014). A tale of an unwanted outcome: Transfers and local endowments of trust and cooperation. *Journal of Economic Behavior & Organization*, 102:74–89.

Agenzia per la Coesione Territoriale (2018). Rapporto sui tempi di attuazione delle opere pubbliche.

Albanese, G., Carrieri, V., Rita Ferrara, A., and Speziali, M. M. (2024). Beyond gdp: assessing the impact of the 2007–13 eu cohesion policy. *Regional Studies*, pages 1–21.

Albanese, G. and De Blasio, G. (2018). European cohesion policy in italy: Empirical evidence and interpretations. Technical report.

Andini, M., Ciani, E., de Blasio, G., D'Ignazio, A., and Salvestrini, V. (2018). Targeting with machine learning: An application to a tax rebate program in italy. *Journal of Economic Behavior & Organization*, 156:86–102.

Antulov-Fantulin, N., Lagravinese, R., and Resce, G. (2021). Predicting bankruptcy of local government: A machine learning approach. *Journal of Economic Behavior & Organization*, 183:681–699.

Aschauer, D. A. (1989). Is public expenditure productive? *Journal of monetary economics*, 23(2):177–200.

Bachtler, J., Berkowitz, P., Hardy, S., and Muravska, T. (2017). *EU Cohesion Policy*. Taylor & Francis.

Baltrunaite, A., Orlando, T., and Rovigatti, G. (2023). The making of public works in italy: Institutional features and regional characteristics. *Italian Economic Journal*, pages 1–38.

Barbero, J., Christensen, M., Conte, A., Lecca, P., Rodríguez-Pose, A., and Salotti, S. (2023). Improving government quality in the regions of the eu and its system-wide benefits for cohesion policy. *JCMS: Journal of Common Market Studies*, 61(1):38–57.

Barone, G., D'Ignazio, A., De Blasio, G., and Naticchioni, P. (2016). Mr. rossi, mr. hu and politics. the role of immigration in shaping natives' voting behavior. *Journal of Public Economics*, 136:1–13.

Becker, S. O., Egger, P. H., and Von Ehrlich, M. (2010). Going nuts: The effect of eu structural funds on regional performance. *Journal of Public Economics*, 94(9-10):578–590.

Becker, S. O., Egger, P. H., and Von Ehrlich, M. (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(4):29–77.

Bloise, F., Brunori, P., and Piraino, P. (2021). Estimating intergenerational income mobility on sub-optimal data: a machine learning approach. *The Journal of Economic Inequality*, 19(4):643–665.

Botchkarev, A. (2018). Performance metrics (error measures) in machine learning regression, forecasting and prognostics: Properties and typology. *arXiv preprint arXiv:1809.03006*.

Bradley, A. P. (1997). The use of the area under the roc curve in the evaluation of machine learning algorithms. *Pattern recognition*, 30(7):1145–1159.

Breiman, L. (2001). Random forests. *Machine learning*, 45:5–32.

Bryk, A. and Schneider, B. (2002). *Trust in schools: A core resource for improvement*. Russell Sage Foundation.

Bucci, M., Ivaldi, G., Messina, G., Moller, L., and Gennari, E. (2021). Infrastructure gaps in italy: A case-by-case measurement. *Bank of Italy Occasional Paper*, (635).

Calderón, C., Moral-Benito, E., and Servén, L. (2015). Is infrastructure capital productive? a dynamic heterogeneous approach. *Journal of Applied Econometrics*, 30(2):177–198.

Caravaggio, N. and Resce, G. (2023). Enhancing healthcare cost forecasting: A machine learning model for resource allocation in heterogeneous regions. *Economics & Statistics Discussion Papers, University of Molise, Department of Economics*.

Carrieri, V., Lagravinese, R., and Resce, G. (2021). Predicting vaccine hesitancy from area-level indicators: A machine learning approach. *Health Economics*, 30(12):3248–3256.

Cawley, G. C. and Talbot, N. L. (2010). On over-fitting in model selection and subsequent selection bias in performance evaluation. *The Journal of Machine Learning Research*, 11:2079–2107.

Cerqua, A., Di Stefano, R., Letta, M., and Miccoli, S. (2021). Local mortality estimates during the covid-19 pandemic in italy. *Journal of Population Economics*, 34(4):1189–1217.

Cerqua, A. and Letta, M. (2022). Local inequalities of the covid-19 crisis. *Regional science and urban economics*, 92:103752.

Cerqua, A. and Pellegrini, G. (2018). Are we spending too much to grow? the case of structural funds. *Journal of Regional Science*, 58(3):535–563.

Cerulli, G. (2023). *Fundamentals of Supervised Machine Learning: With Applications in Python, R, and Stata*. Springer Nature.

Charbuty, B. and Abdulazeez, A. (2021). Classification based on decision tree algorithm for machine learning. *Journal of Applied Science and Technology Trends*, 2(01):20–28.

Chaurey, R. and Le, D. T. (2022). Infrastructure maintenance and rural economic activity: Evidence from india. *Journal of Public Economics*, 214:104725.

Chicco, D., Warrens, M. J., and Jurman, G. (2021). The coefficient of determination r-squared is more informative than smape, mae, mape, mse and rmse in regression analysis evaluation. *Peerj computer science*, 7:e623.

Coco, G. and Lagravinese, R. (2014). Cronyism and education performance. *Economic Modelling*, 38:443–450.

Coco, G. and Lagravinese, R. (2021). Cohesion policy and public investment in the eu. In Cerniglia, F., Saraceno, F., and Watt, A., editors, *The great reset: 2021 European public investment outlook*, page 203. Open Book Publishers.

Coco, G., Lagravinese, R., and Resce, G. (2020). Beyond the weights: a multicriteria approach to evaluate inequality in education. *The Journal of Economic Inequality*, 18:469–489.

Coco, G., Monturano, G., and Resce, G. (2023). Assessing the effectiveness of coordination among public authorities in cohesion expenditure. In *BOOK OF SHORT PAPERS*, pages 329–334.

Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and psychological measurement*, 20(1):37–46.

Coppola, G., Destefanis, S., Marinuzzi, G., and Tortorella, W. (2018). European union and nationally based cohesion policies in the italian regions. *Regional Studies*.

Crescenzi, R. and Giua, M. (2020). One or many cohesion policies of the european union? on the differential economic impacts of cohesion policy across member states. *Regional Studies*, 54(1):10–20.

Cunico, G., Aivazidou, E., and Mollona, E. (2022). Decision-making traps behind low regional absorption of cohesion policy funds. *European Policy Analysis*, 8(4):439–466.

Davoudi, S. and Strange, I. (2009). *Conceptions of space and place in strategic spatial planning*. Routledge London.

De Blasio, G., D'Ignazio, A., and Letta, M. (2022). Gotham city. predicting 'corrupted'municipalities with machine learning. *Technological Forecasting and Social Change*, 184:122016.

De Mol, C., De Vito, E., and Rosasco, L. (2009). Elastic-net regularization in learning theory. *Journal of Complexity*, 25(2):201–230.

Delogu, M., Lagravinese, R., Paolini, D., and Resce, G. (2024). Predicting dropout from higher education: Evidence from italy. *Economic Modelling*, 130:106583.

DeLong, E. R., DeLong, D. M., and Clarke-Pearson, D. L. (1988). Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics*, pages 837–845.

Di Stefano, R. and Resce, G. (2024). The determinants of missed funding: Predicting the paradox of increased need and reduced allocation. *Available at SSRN 4663110*.

D'Alpaos, C., Moretto, M., Valbonesi, P., and Vergalli, S. (2013). Time overruns as opportunistic behavior in public procurement. *Journal of Economics*, 110:25–43.

Egwim, C. N., Alaka, H., Toriola-Coker, L. O., Balogun, H., and Sunmola, F. (2021). Applied artificial intelligence for predicting construction projects delay. *Machine Learning with Applications*, 6:100166.

Espinoza, R. and Presbitero, A. F. (2022). Delays in public investment projects. *International Economics*, 172:297–310.

European Commission (2022). *8th Cohesion Report: Cohesion in Europe towards 2050*. Luxembourg: Publications Office of the European Union, 2022.

Fawcett, T. (2006). An introduction to roc analysis. *Pattern recognition letters*, 27(8):861–874.

Ferry, M. and McMaster, I. (2013). Cohesion policy and the evolution of regional policy in central and eastern europe. *Europe-asia studies*, 65(8):1502–1528.

Fratesi, U. and Perucca, G. (2018). Territorial capital and the resilience of european regions. *The Annals of Regional Science*, 60(2):241–264.

Friedman, J. H. (2001). Greedy function approximation: a gradient boosting machine. *Annals of statistics*, pages 1189–1232.

Gondia, A., Siam, A., El-Dakhakhni, W., and Nassar, A. H. (2020). Machine learning algorithms for construction projects delay risk prediction. *Journal of Construction Engineering and Management*, 146(1):04019085.

Gori, G. F., Lattarulo, P., and Mariani, M. (2017). Understanding the procurement performance of local governments: A duration analysis of public works. *Environment and planning C: politics and space*, 35(5):809–827.

Gori, G. F., Lattarulo, P., Mariani, M., and Razzolini, L. (2024). The expediting effect of monitoring on infrastructural works. a regression-discontinuity approach with multiple assignment variables. *Italian Economic Journal*, 10(1):197–224.

Greenwell, B. M. et al. (2017). pdp: An r package for constructing partial dependence plots. *R J.*, 9(1):421.

Guzzardi, D., Palagi, E., Roventini, A., and Santoro, A. (2023). Reconstructing income inequality in italy: New evidence and tax system implications from distributional national accounts. *Journal of the European Economic Association*.

Hand, D. J. and Till, R. J. (2001). A simple generalisation of the area under the roc curve for multiple class classification problems. *Machine learning*, 45:171–186.

He, C., Milne, A., and Ataullah, A. (2023). What explains delays in public procurement decisions? *Economic Modelling*, 121:106201.

Hearst, M. A., Dumais, S. T., Osuna, E., Platt, J., and Scholkopf, B. (1998). Support vector machines. *IEEE Intelligent Systems and their applications*, 13(4):18–28.

Incaltarau, C., Pascariu, G. C., and Surubaru, N.-C. (2020). Evaluating the determinants of eu funds absorption across old and new member states–the role of administrative capacity and political governance. *JCMS: Journal of Common Market Studies*, 58(4):941–961.

Jean, N., Burke, M., Xie, M., Davis, W. M., Lobell, D. B., and Ermon, S. (2016). Combining satellite imagery and machine learning to predict poverty. *Science*, 353(6301):790–794.

Khwaja, A. I. (2009). Can good projects succeed in bad communities? *Journal of public Economics*, 93(7-8):899–916.

Kleinberg, J., Ludwig, J., Mullainathan, S., and Obermeyer, Z. (2015). Prediction policy problems. *American Economic Review*, 105(5):491–495.

Kramer, O. and Kramer, O. (2013). K-nearest neighbors. *Dimensionality reduction with unsupervised nearest neighbors*, pages 13–23.

Krstajic, D., Buturovic, L. J., Leahy, D. E., and Thomas, S. (2014). Cross-validation pitfalls when selecting and assessing regression and classification models. *Journal of cheminformatics*, 6:1–15.

Kuhn, M. (2015). Caret: classification and regression training. *Astrophysics Source Code Library*, pages ascl–1505.

Kuhn, M., Wing, J., Weston, S., Williams, A., Keefer, C., Engelhardt, A., Cooper, T., Mayer, Z., Kenkel, B., Team, R. C., et al. (2020). Package 'caret'. *The R Journal*, 223(7).

Lagravinese, R. (2015). Economic crisis and rising gaps north–south: evidence from the italian regions. *Cambridge journal of regions, economy and society*, 8(2):331–342.

Leduc, S. and Wilson, D. (2013). Roads to prosperity or bridges to nowhere? theory and evidence on the impact of public infrastructure investment. *NBER Macroeconomics Annual*, 27(1):89–142.

Liu, Y., Wang, Y., and Zhang, J. (2012). New machine learning algorithm: Random forest. In *Information Computing and Applications: Third International Conference, ICICA 2012, Chengde, China, September 14-16, 2012. Proceedings 3*, pages 246–252. Springer.

McHugh, M. L. (2012). Interrater reliability: the kappa statistic. *Biochemia medica*, 22(3):276–282.

Ministero per lo Sviluppo Economico (2007). Quadro strategico nazionale per la politica regionale di sviluppo 2007-2013. Technical report.

Murphy, K. P. et al. (2006). Naive bayes classifiers. *University of British Columbia*, 18(60):1–8.

Natekin, A. and Knoll, A. (2013). Gradient boosting machines, a tutorial. *Frontiers in neurorobotics*, 7:21.

Neilson, C. A. and Zimmerman, S. D. (2014). The effect of school construction on test scores, school enrollment, and home prices. *Journal of Public Economics*, 120:18–31.

Nifo, A. and Vecchione, G. (2014). Do institutions play a role in skilled migration? the case of italy. *Regional Studies*, 48(10):1628–1649.

Palcic, D., Reeves, E., Flannery, D., and Geddes, R. R. (2022). Public-private partnership tendering periods: an international comparative analysis. *Journal of Economic Policy Reform*, 25(2):156–172.

Peterson, L. E. (2009). K-nearest neighbor. *Scholarpedia*, 4(2):1883.

Ramey, V. A. et al. (2021). *The macroeconomic consequences of infrastructure investment*, volume 28215. University of Chicago Press Chicago.

Reeves, E., Palcic, D., Flannery, D., and Geddes, R. R. (2017). The determinants of tendering periods for ppp procurement in the uk: an empirical analysis. *Applied economics*, 49(11):1071–1082.

Resce, G. and Vaquero-Piñeiro, C. (2022). Predicting agri-food quality across space: A machine learning model for the acknowledgment of geographical indications. *Food Policy*, 112:102345.

Rish, I. et al. (2001). An empirical study of the naive bayes classifier. In *IJCAI 2001 workshop on empirical methods in artificial intelligence*, volume 3, pages 41–46. Citeseer.

RodrÍguez-Pose, A. and Fratesi, U. (2004). Between development and social policies: the impact of european structural funds in objective 1 regions. *Regional studies*, 38(1):97–113.

Rodríguez-Pose, A. and Garcilazo, E. (2015). Quality of government and the returns of investment: Examining the impact of cohesion expenditure in european regions. *Regional Studies*, 49(8):1274–1290.

Sansone, D. (2019). Beyond early warning indicators: High school dropout and machine learning. *Oxford bulletin of economics and statistics*, 81(2):456–485.

Sansone, D. and Zhu, A. (2023). Using machine learning to create an early warning system for welfare recipients. *Oxford Bulletin of Economics and Statistics*, 85(5):959–992.

Schmidhuber, J. (2015). Deep learning in neural networks: An overview. *Neural networks*, 61:85–117.

Steinwart, I. and Christmann, A. (2008). *Support vector machines*. Springer Science & Business Media.

Surubaru, N.-C. (2017). Administrative capacity or quality of political governance? eu cohesion policy in the new europe, 2007–13. *Regional Studies*, 51(6):844–856.

Sutskever, I., Vinyals, O., and Le, Q. V. (2014). Sequence to sequence learning with neural networks. *Advances in neural information processing systems*, 27.

Tharwat, A. (2021). Classification assessment methods. *Applied computing and informatics*, 17(1):168–192.

Titl, V., Mazrekaj, D., and Schiltz, F. (2024). Identifying politically connected firms: A machine learning approach. *Oxford Bulletin of Economics and Statistics*, 86(1):137–155.

Tosun, J. (2014). Absorption of regional funds: A comparative analysis. *JCMS: Journal of Common Market Studies*, 52(2):371–387.

Wong, H. L., Wang, Y., Luo, R., Zhang, L., and Rozelle, S. (2017). Local governance and the quality of local infrastructure: Evidence from village road projects in rural china. *Journal of Public Economics*, 152:119–132.

Zou, H. and Hastie, T. (2005). Regularization and variable selection via the elastic net. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 67(2):301–320.

Chapter 3

Influence of Poverty and Inequalities on the Italian Minimum Income Guarantee

An Analysis with Spatially Clustered Regression*

Abstract

We investigate the effects of poverty and inequality on the distribution of the Italian Citizenship Income policy, which aims to support financially disadvantaged households. Using spatial econometric models, we analyse how income support correlates with local wellbeing indicators such as per capita income, poverty rates, and the Gini index. These models help us understand the policy's coverage by examining the spatial heterogeneity of recipient households and their local economic contexts. Our findings indicate that poverty and inequality significantly influence the geographical distribution of policy participation. In particular, areas with high socioeconomic deprivation and low income levels show a positive correlation with policy reach. In contrast, regions with higher incomes and greater inequality see less policy engagement. The analysis highlights the complex socioeconomic landscape in Italy, further complicated by the COVID-19 pandemic's impact on regional disparities.

JEL classifications: H53, I38, R12, C21

Keywords: Guaranteed Minimum Income, Municipal Income Inequality, Spatially Clustered Regression, Policy Evaluation, Spatial Heterogeneity and Complexity

---

*This work is paper co-authored with **Dr. Paolo Maranzano** (Department of Economics, Management and Statistics, University of Milano-Bicocca, Piazza dell'Ateneo Nuovo, 1, Milan, 20126, Italy) and **Professor Pasquale Tridico** (Department of Economics, Roma Tre University, Via Silvio D'Amico 77, Rome, 00145, Italy), and is published in Working Paper version in researchsquare.com https://www.researchsquare.com/article/rs-4271005/v1. We thank the Italian Institute of Social Security (INPS) for providing the data.

3.1 Introduction

In recent decades, due to the recessive effects of several economic crises generated on a global scale, there has been a rapid increase in poverty and inequalities, which affect the socioeconomic sphere and the wellbeing of citizens (Jenkins et al., 2012; Stiglitz, 2012; Piketty and Saez, 2014; Aaberge and Brandolini, 2015; Fadda and Tridico, 2017).

The acceleration of these phenomena, as shown in a large number of studies, might depend on multiple interconnected factors and variables, such as the automation of production processes with the transition from human labour to technology (Acemoglu and Restrepo, 2018, 2022; Gregory et al., 2022; Moll et al., 2022), globalisation (Milanovic, 2016; Ravallion, 2018; Nolan et al., 2019), intergenerational mobility (Raitano and Vona, 2015; Bavaro and Patriarca, 2022), the flexibilisation of the labour market (Tridico, 2018), up to differentiated access to education systems (Franzini et al., 2020; Bonacini et al., 2021).

Today, the income gap between different segments of the world population is increasingly marked, especially considering specific and distinct territorial contexts (Rodríguez-Pose and Hardy, 2015). From an empirical perspective, a substantial volume of scientific studies, such as the works of Alesina et al. (2004), Checchi and Peragine (2010), Iammarino et al. (2019), Chancel and Piketty (2021), Pignataro (2021), Liberati and Resce (2022), have highlighted the decisive and central role of spatial heterogeneity, and geographical distance, in explaining income differentials.

In spatial terms, inequality can be classified along two different dimensions depending on how it interacts with space: a "within" dimension concerning disparities within the same country or community, and a "between" dimension which refers, instead, to the differences across different states, regions and territories (Bourguignon and Morrisson, 2002; Liberati, 2015; Chancel and Piketty, 2021). From this classification, we can observe a greater concentration of inequality in highly urbanised areas, with a growing dualism between suburbs, characterised by greater social and material vulnerability, and centres (Lelo et al., 2019; Nijman and Wei, 2020). Unlike in less densely populated areas, such as rural villages and remote communities, fewer internal disparities are observed, due to lower and more homogeneous income levels on average, determined by limited job opportunities, mainly concentrated in economic sectors with low added value and with limited degrees of technical innovation, such as agriculture and small local and family businesses (Rodríguez-Pose and Hardy, 2015). However, taking into account the distance between areas (inequality between), it can be observed that lower rurality leads to lower income inequalities (Zhong et al., 2022). Poverty also appears to follow spatial patterns similar to those of socioeconomic disparities, with a greater concentration of situations of social exclusion and vulnerability in the most remote areas and places or in the extreme outskirts of cities with high population density (Rodríguez-Pose and Hardy, 2015; Zhong et al., 2022).

To try to limit the ubiquity of these phenomena, political decision makers, at all levels, are increasingly committing important public resources to the organisation of coherent welfare programmes and interventions, aimed at supporting citizens in difficult situations (Atkinson and Bourguignon, 2014; Atkinson, 2015; Blanchard and Rodrik, 2023). In fact, over the years, many international institutions, and in particular the European Union, national states and regions have oriented their policies towards the activation of social protection schemes against poverty and inequalities. Currently, political attention to the issue of mitigation has been further accentuated by the COVID-19 pandemic, which has widened the disparity between citizens, worsening pre-existing situations of poverty and inequality (Palomino et al., 2020; Cerqua and Letta, 2022; Gallo and Raitano, 2023). To a large extent, these programs have been organized in the form of subsidies and direct transfers or through structured guaranteed minimum income schemes, based on specific conditions, for example, the level of family or individual income (Immervoll and Scarpetta, 2012).

However, despite the increased institutional attention to these issues, the effectiveness of the measures adopted by governments to support income is still much debated at the level of scientific literature and there is no univocal consensus, especially if the distribution of take-up rate is taken into account[1] of minimum income schemes (Bramley et al., 2000; Bargain et al., 2012; Bhargava and Manoli, 2015; Aprea et al., 2024; Boscolo and Gallo, 2024).

3.1.1 Our Contribution

This article investigates the influence of income inequality and poverty on a specific Italian public policy of guaranteed minimum income, intended for families and individuals in difficulty, namely the *Reddito di Cittadinanza* (in English, *Citizenship Income*, hereinafter referred to as Guaranteed Minimum Income or GMI), approved in 2019 and implemented until 2023. The main goal of the GMI policy was to support the income of Italian families living below the poverty threshold, through the provision of a guaranteed minimum income, conditioned on specific eligibility criteria. However, during the entire implementation phase, the GMI was subject to multiple criticisms and public and political discussions, pertaining to cases of fraudulent claims and receipts of the benefit, the possible employability of the recipients, and the coverage of the eligible target units (Busilacchi and Fabbri, 2023; Tonutti et al., 2022; Maitino et al., 2024).

[1]In economics, take-up refers to the rate of enrolment in a social programme or benefit, such as subsidies or direct transfers against poverty. It measures the percentage of eligible people who actually apply for or are assisted by the benefit. In other words, it represents the proportion of subjects eligible for a given policy who manage to benefit from it. A low acceptance rate can be caused by several factors, such as inadequate information or communication, technical bureaucratic obstacles, or social barriers related to the condition of applicants.

In general, the economic literature, as described in Section 3.2, shows that income support measures generate heterogeneous results, mainly deriving from the nature of the measures themselves and their access conditions (e.g., families can require support only if below a specific income), as well as from the time frame and territorial characteristics in which they operate (Almeida et al., 2022; Guimarães and Lourenço, 2024). Motivated by such a general uncertainty on the effectiveness of the programmes, as well as by the high amount of public resources invested in the period of operation of the GMI measure, that is, around 23 billion euros between 2019 and 2022 (INPS, 2022a,b), we stress the importance of analysing the socioeconomic determinants of geographical heterogeneity in the rate of access of Italian families to the GMI program. In particular, following the *take-up* framework on minimum income measures (Bargain et al., 2012; Bhargava and Manoli, 2015; Aprea et al., 2024; Boscolo and Gallo, 2024) and largely on the literature focusing on the *participation rates* in the policy programmes (Hernandez et al., 2007; Åslund and Fredriksson, 2009; Chetty et al., 2013; Markussen and Røed, 2015; Bauer and Dang, 2016; Grossman and Khalil, 2020), we are interested in studying how the Italian income support programme (i.e., the proportion of families requiring the GMI in a certain year) varied across the Italian municipalities according to local socioeconomic and wellbeing determinants, that is, the average per capita income, the municipal share of poverty, and the Gini index on the declared income (these indicators are illustrated in section 3.3.2). In synthesis, our study aims to explore the distribution of the number of recipients' families across the Italian municipalities concerning two dimensions: 1) geography and spatial heterogeneity; 2) the role of poverty and inequality in determining GMI participation.

Regarding the geographical dimension, many social and economic statistics at the territorial level show that Italy is strongly characterised by heterogeneity across spatial units and areas, i.e., there exists an apparent spatial heterogeneity as defined by the Second Law of Geography (Zhu and Turner, 2022). In fact, within the western bloc and, in particular, compared to other European countries, Italy presents high levels of socioeconomic and income inequality, with considerable heterogeneity within the country (Simonazzi et al., 2013; Celi et al., 2017; Tridico, 2018). This internal heterogeneity, clearly illustrated in Figure 3.2, derives from the historical divergence in growth and economic development between the rich regions of the North and the poorer regions of the South, which today are becoming increasingly marked (Salvati and Carlucci, 2014; Mussida and Parisi, 2019). Such a marked spatial heterogeneity weakens the implementation and the efficiency of public policies, in particular, those related to social assistance and protection measures (Albanese et al., 2023), in the Italian context. Looking at the municipal level data represented in Figure 3.1, the distribution of the GMI was also conditioned by the typical Italian dualism. In fact, a much higher concentration of households was recorded in the southern regions, compared to the central and northern regions, during the entire period of the activity of the measure (Busilacchi and Fabbri, 2023; Monturano et al., 2023b; Maitino et al., 2024).

We address this issue by adopting a spatial econometric methodology, namely the

Spatially-Clustered Regression (SCR) approach of (Sugasawa and Murakami, 2021), which breaks up the entire national territory by clustering municipal units into homogeneous and spatially contiguous groups while estimating local relationships via regression. In particular, the method is used to obtain global and local estimates of the influence of several socioeconomic and administrative factors, including poverty and inequality, on the number of households requiring and obtaining the GMI income support measure at the most granular level possible for Italy. Furthermore, to investigate the temporal dynamics of the relationship between inequality, poverty, and participation in income support, the spatial regression algorithm is replicated for municipal data from 2019 to 2022, thus fully including the peak period of the COVID-19 pandemic. This provides several policy implications, including whether and how the relationships have changed during the political period and whether these changes affected the entire country or only some areas.

To the best of our knowledge, this work represents the first attempt conducted at a local granular level to investigate the influence of socioeconomic disparities on the GMI, jointly exploiting variables that change across space and time. We stress that we do not seek to estimate a causal relationship among socioeconomic variables and the coverage (in terms of reached families) of GMI policy. Our primary goal is to identify statistically significant relationships among the above-mentioned variables, expected to be non-constant across space and time due to the spatial heterogeneity and the temporal dynamic. Specifically, we are interested in evaluating how these relationships vary across the Italian areas and how the role of poverty and inequality changed according to different geographical and administrative partitions of Italy. The results highlight how the implementation of the minimum income measure is strictly interconnected to Italian socio-spatial dynamics. In particular, we detect a non-homogeneous, but nevertheless growing, link between levels of income inequality and recipient households over the time the policy was operational. The spatial correlation, however, is stronger with regard to the distribution of per capita incomes and the share of municipal poverty. In general, these results underline the importance of organising and adapting public minimum income policies concerning different local specificity and characteristics, in order to ensure that these monetary transfers properly reach the target families.

The remainder of the paper is structured as follows. Section 3.2 presents a review of the literature on guaranteed minimum income and the description of the Italian GMI. Section 3.3 examines the available data at the municipal level for Italy and provides exploratory descriptive statistics. Section 3.4 describes the spatially-clustered regression methodology used to estimate the global and local effect of socioeconomic weaknesses on GMI participation. Section 3.5 presents the empirical strategy used for the estimates. In particular, we provide the econometric specifications used in the analysis and the rationale behind the empirical choices. Section 3.6 summarises and discusses the main findings obtained. Policy implications are also discussed. Finally, in Section 3.7, we synthesise the paper's contents and provide conclusive remarks.

3.2 Guaranteed Minimum Income: Theory and Application in Italy

Although the protection measures developed through guaranteed minimum income schemes have a consolidated history in relation to the fight against poverty and the mitigation of the effects of unemployment (Marx and Nelson, 2013; Baldini et al., 2018a; Natili, 2020; Busilacchi and Fabbri, 2023), only recently, due to the increase in economic disparities, have these issues been assigned an importance, on an international scale, as effective measures for combating socioeconomic inequalities (Baldini et al., 2018b; Gallo, 2021). Historically, in fact, minimum income policies have evolved considerably compared to current formulations. As highlighted in various works, such as Marx and Nelson (2013), Baldini et al. (2018a), Natili (2020), these policies have undergone different phases of transformation, driven mainly by political and institutional concerns related to the efficiency of benefits in relation to their capacity to promote active inclusion. Initially, in fact, until the 1980s, they were conceived as protection tools, to prevent situations of vulnerability by providing financial support to individuals in conditions of serious poverty (Marx and Nelson, 2013; Natili, 2020; Busilacchi and Fabbri, 2023). Subsequently, to try to avoid the risk of a so-called "welfare trap", i.e. the situation in which beneficiaries become dependent on state assistance, distancing themselves from the labour market (Immervoll et al., 2015), minimum income policies have embraced the paradigm of work activation (Hemerijck, 2017; Natili, 2020), and the concept of workfare (Xu and Carraro, 2017; Groot et al., 2019; Spies-Butcher, 2020). These developments have profoundly transformed the initial objectives of these policies, which have evolved from simple passive income support mechanisms to dynamic tools aimed at encouraging the work and social integration of the beneficiaries. Currently, the principles of activation and workfare are changing further, away from a universalistic perspective of the benefit (Van Parijs and Vanderborght, 2017; Groot et al., 2019; Spies-Butcher, 2020). In many countries, especially on a micro-territorial scale, forms of universal minimum income (UBI) (Groot et al., 2019; Feinberg and Kuehn, 2020; Banerjee et al., 2023) are being tested, focused simply on residence in a specific place, therefore, without any requirement for proof of income. Yet, in this framework of change, which expresses the political will to make minimum income schemes more efficient and standardised, observation of the various measures approved on an international scale reveals a still very fragmented and heterogeneous picture, in terms of structure, implementation and effectiveness (Natili, 2020; Aprea et al., 2022).

At the scientific literature level, following Immervoll and Scarpetta (2012), we can identify the *guaranteed minimum income* as a plurality of welfare and social assistance interventions usually aimed at families in poverty, connected to the verification of economic means. Interventions aimed at integrating income from work or pension, education, housing emergency, employment, disability, etc. can, therefore, also be considered forms of minimum income. In practical terms, these measures represent direct monetary transfers or services useful to supplement the incomes of low-income families.

Generally, especially in European countries, they are organised on a principle of selective universalism, as allows for: 1) examination of family incomes globally, without considering specific groups of individuals or social classes; 2) present specific eligibility criteria and conditions (Natili, 2020).

Frazer and Marlier (2016), Curci et al. (2020) and Aprea et al. (2022) highlight how the thresholds for access to minimum income schemes are often represented by multiple socioeconomic and/or wellbeing indicators, such as a specific poverty threshold, or by the minimum wage level or unemployment benefits. In this regard, the European Union (EU) anchors the receipt of these measures at a specific relative poverty threshold, defined as AROPE[2], below which one is considered poor in the EU. This value represents 60% of the national median equivalent disposable income (Atkinson et al., 2015). Despite this, minimum income schemes are often subordinated, not only to a calculated level of deprivation or vulnerability, but also to other factors, such as citizenship and/or long-term residence requirements, family size and structure, as well as other regulations related to wealth levels and participation in inclusion and training programmes. These differences contribute to increasing the heterogeneity between the various regulatory systems pertaining to these policies.

At community level, the European institutions have repeatedly requested the introduction of specific guaranteed minimum income schemes in all member states (Natili, 2020). Today, minimum income schemes are present in all countries of the Union. Even with differences in terms of time extension, amounts, and the number of eligible people, they present many similarities concerning the conditions for accessing the benefit.

However, until 2017, Italy and Greece, the European countries most marked by internal disparities, and among the most impacted by the increase in inequalities and large-scale social exclusion, were the only two EU member states not to have a national programme including income support (Gallo, 2021; Busilacchi and Fabbri, 2023; Maitino et al., 2024). Only in 2018 has Italy adopted a national measure for income support with the Inclusion Income (ReI)[3]. Minor protection schemes, lacking universality in treatment were, however, already present at a national level[4] and in various regions; for a review of these measures see Gallo (2021) and Busilacchi and Fabbri (2023). The Inclusion Income has had a very short time of just over one year since its inception, after which it was replaced by the Italian

[2]Eurostat defines the AROPE rate ("at risk of poverty or social exclusion") as the sum of people who are at risk of poverty, or who are seriously deprived materially and socially or who live in a family with a very low work intensity.

[3]The Inclusion Income ReI, Law 15 March 2017, n. 33, is a universal measure to combat poverty, conditioned on the evaluation of the economic condition. It was abolished in 2019 and replaced with Citizenship Income, GMI.

[4]The Inclusion Income (ReI) abolished the Support for Active Inclusion (SIA) and the Unemployment Allowance (ASDI).

Citizenship Income (GMI); a more structured guaranteed minimum income policy which was allocated three times more public resources than had been previously allocated (see the section below, Section 3.2.1, for a detailed overview of the measure).

Generally speaking, although guaranteed minimum income measures have a consolidated history, the effects of mitigating poverty and inequalities are not yet entirely clear (Marx and Nelson, 2013). In fact, we can observe differentiated impacts in terms of direction and intensity, both in advanced countries and in poorer states. For example, in China (Yu and Li, 2021) and other Asian states (Wagle, 2017), despite the pronounced regional differences, a positive and lasting correlation is found between social security spending and urban-rural income disparity, particularly in Asian states with lower incomes. Within the African continent, government transfers appear to be more correlated with income growth and long term economic expansion and employment (Maket et al., 2023). In fact, Banerjee et al. (2023) presents an extensive empirical investigation, conducted with data obtained in the field in remote villages of Kenya, and highlights how the provision of a universalistic minimum income and other forms of income support aimed at poor local populations, have led to changes in occupational choices, with the transition from employment to entrepreneurship, as well as significant economic expansion, with important implications on saving, access to credit, and investment behaviour. In contrast, in western countries, which allocate larger shares of GDP to assistance, these programmes seem to have a more nuanced impact (Almeida et al., 2022; Rauh and Santos, 2022; Gallo and Raitano, 2023; Guimarães and Lourenço, 2024), if not potentially negative in terms of macroeconomic evaluations (Conesa et al., 2023; Connolly et al., 2024; Daruich and Fernández, 2024). Despite this, even in these richer countries, there are positive effects of these measures on income levels and poverty, with specific reference to homeless citizens (Gubits et al., 2018; Locks and Thuilliez, 2023), on unemployment (Terracol, 2009; Card et al., 2015, 2018; Calnitsky, 2020), and wage inequality (Calnitsky, 2020; Cantillon et al., 2020).

In the Italian context, despite great initial expectations, the GMI, throughout its period of activity, has been accompanied by numerous critical issues and discussions. These criticisms have been mainly linked to cases of fraud in obtaining the benefit (Monturano, 2023). Other points debated have concerned the effectiveness of the GMI in making recipients active and employable (Busilacchi and Fabbri, 2023; Maitino et al., 2024), the take-up rates connected to the income criteria of the measure (Ansaloni et al., 2024; Boscolo and Gallo, 2024), as well as the targeting capacity of the GMI, considered as a sort of generalised subsidy for the southern regions (Monturano et al., 2023b). Some studies have specifically focused on assessing the impacts of the GMI on poverty, with quantitative estimates of small areas (Tonutti et al., 2022), and inequalities, through fiscal microsimulation estimates (Gallo, 2021), with divergent results.

Within this literature, at least to our knowledge, no work has so far studied the influence of income inequality and poverty on the measures of the Italian minimum income while also considering the spatial component. In this sense, given the spatial heterogeneity present in Italy, the Italian Citizenship Income emerges as a unique experiment to study the effectiveness of these policies.

3.2.1 The Italian GMI: Characteristics and Prerequisites

The Italian Citizenship Income (in Italian, Reddito di Cittadinanza)[5] represented the most structured policy of guaranteed minimum income, aimed at families and individuals to combat poverty and inequalities approved in Italy. Despite the name, which refers to universal basic income schemes, this income support is structured as a guaranteed minimum income, focused on monetary transfers to low-income families. The measure also included active employment policies, aimed at encouraging employment and social inclusion, with a view to workfare.

The programme was structured on the principle of selective universality, conditioned by the verification of economic resources, with the Equivalent Economic Situation Indicator (ISEE)[6]. Along with the income requirement, the benefit was subject to a number of other criteria, such as possession of Italian citizenship, or continuous residence for at least ten years, defined family structures and specific sizes, certain real estate and financial assets, etc. In this sense, access to the benefit was not guaranteed by simply meeting the requirements of a single specific criterion, because to be admitted to the programme it was necessary to simultaneously satisfy all of the eligibility conditions (see Table C1 for the complete details of the access requirements to the GMI. The benefit had a renewable duration of 18 months. The amount varied from 780 to 1716 euros per month, depending on the family composition and any rent costs.

During its implementation phase, the GMI reached a very large number of families. Before the proliferation of the COVID-19 pandemic, the measure had already reached over 1 million families. These figures grew dramatically during the two-year period 2020-2021, the period most affected by the lockdown measures (INPS, 2023). In the last year of activation, 2023, the number of beneficiaries decreased, due to the change in the GMI regulations, which reduced the number of eligible people. These numbers reflect the importance that the GMI has assumed throughout the health emergency in limiting the spread of poverty and inequalities among Italian families (INPS, 2022b,a; Gallo and Raitano, 2023).

A different trend is recorded in relation to the average regional amounts per unit, which have shown an ever increasing trend along the entire series of the operations of the measure, rising from 492.17 Euros in 2019 to 562.81 Euros in 2023[7].

[5]The Citizenship Income was introduced with Legislative Decree no. 4 of 28 January 2019, converted with Law no. 26 of 28 March 2019.

[6]The Equivalent Economic Situation Indicator (ISEE) for obtaining the GMI referred to the year preceding the request.

[7]The data is present at this link: https://www.inps.it/it/it/dati-e-bilanci/osservatori-statistici-e-altre-statistiche/dati-cartacei---rdc.html.

Figure 3.1 shows the spatial distribution relating to the share of GMI beneficiaries at a local level (Table C1 instead contains the aggregate values on a regional and macro-regional scale). The GMI values refer to the years between 2019 and 2022, i.e. the entire time interval in which the policy operated without any regulatory changes in the eligible population[8].

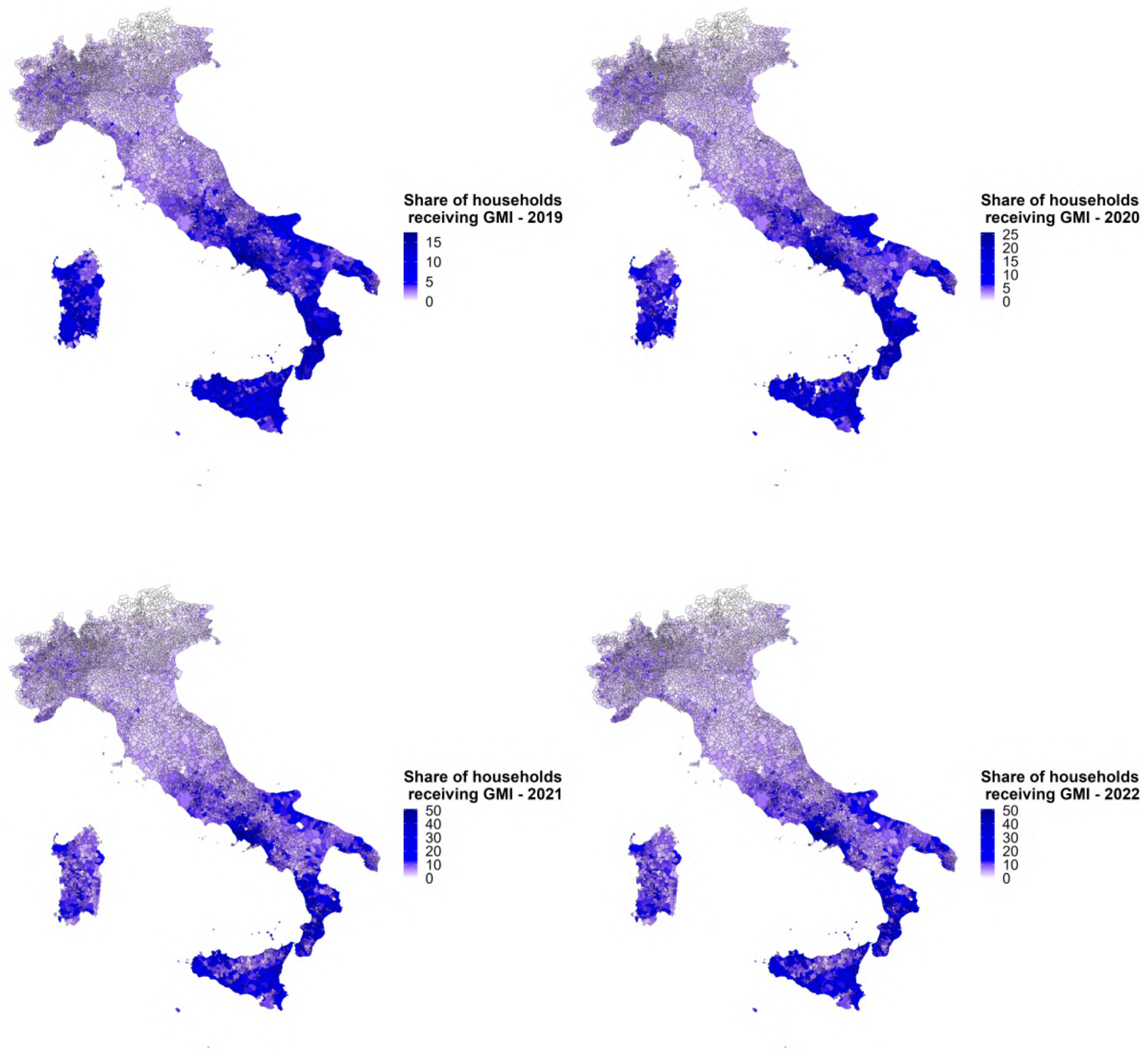
Figure 3.2 instead reports the average municipal values of three socioeconomic variables at the local level, namely income inequality (calculated using the Gini index), average per capita income, and the proportion of citizens in poverty, for the years 2018-2021, therefore, starting from the period immediately preceding the approval of the GMI (the Table C2 presents statistics on these variables at more aggregated territorial levels).

At a comparative level, Figures 3.1 and 3.2 offer useful insights for studying the geographical distribution of economic disparities in Italy and of families who received Citizenship Income. From these, we can draw some interesting observations, highlighting direct socio-spatial connections between inequalities and poverty and obtaining the monetary benefits of the GMI. In general, it is observed that the areas characterised by low levels of per capita income, lower inequalities, and higher shares of poverty are those with a greater number of families benefiting from the GMI. In fact, there is a greater concentration of beneficiaries in the regions of Southern Italy and large cities. For example, regions such as Calabria, Campania, and Sicily jointly present high proportions of poverty (in the South and the Islands, the poverty proportion is above 40%, see Table C2), a high percentage of families treated (about 15%, relative to the ratio between families receiving GMI and the total number of families in these regions, Table C1), as well as per capita incomes and lower economic inequalities. In contrast, in territorial contexts characterised by higher rates of wellbeing and a more prosperous economy, such as regions of Northern and Central Italy (Mussida and Parisi, 2019), there is less need for welfare interventions. Regions such as Lombardy, Emilia-Romagna, Veneto, and Trentino Alto Adige have lower poverty proportions. Therefore, the registered share of families benefiting from GMI is much lower than that of the South (about 2% in the North-East in 2022, see Table C1).

These findings suggest a strong link between the need for state assistance and places where socioeconomic indicators reflect lower levels of wellbeing, in terms of poverty and social exclusion. In particular, these trends seem to highlight the fact that Citizenship Income has played a crucial role in supporting the needs of families in difficulty, particularly in more economically and socially deprived areas but explains less in terms of fighting inequalities of income. However, to fully understand the nature and influence of these links and their implications regarding public income support policies, and in particular, to evaluate the GMI, it is necessary to conduct an in-depth analysis, which also takes into consideration other different socioeconomic, geo-demographic, political-institutional and labour market variables. In this sense, spatial econometric analysis becomes useful to understand the basic relationships between these phenomena.

[8] The GMI was abolished from 1 January 2024, with the Law 29 December 2022, No 197.

Figure 3.1: Spatial Distribution of Average Municipal Shares of Households Receiving the GMI of Total Households (2019-2022)



Source: Author's processing of INPS data.

3.3 Dataset and Descriptive Statistics

3.3.1 Data and Socio-economic Variables

Starting from studies conducted on micro-territorial data, for example, Antulov-Fantulin et al. (2021), Resce (2022), Monturano et al. (2023a), we build a rich panel dataset on municipal variables, collected from multiple national statistical sources[9]. The available data are capable of identifying economic, social, demographic, geographical, environmental, and administrative-institutional conditions that could have influenced the choices of households regarding participation in the GMI programme (see Table for the complete details of the variables used in Table C1). In particular, following ISTAT (2024), we consider as relevant and predominant in relation to the others, three socioeconomic variables, calculated by starting with the income data provided by the Italian Ministry of Economy and Finance, that are capable of identifying the social and economic wellbeing at the local level[10], that is:

- *Income inequality*: the Gini index on declared income obtained by aggregating data from different municipal income groups. Following Antulov-Fantulin et al. (2021), Resce (2022), the Gini index for income inequality is calculated by:

$$G = \frac{\sum_{j=1}^{k}(F_j - F_{j-1}) \times (A_j + A_{j-1})}{2 \times N \times \sum_{j=1}^{k} a_j} \tag{3.1}$$

  where:
  - $F_j = \sum_{i=1}^{j} f_i$ represents the cumulative frequency of individuals declaring an income up to the j-th ordered income bracket (with f_i being the share of the population over the total associated with the i-th ordered income class);
  - F_{j-1} represents the cumulative frequency of declaring an income up to the previously ordered income bracket;
  - A_j represents the cumulative amount of income declared up to the j-th ordered income bracket (with a_i being the share of total declared income associated with the i-th ordered income class);
  - A_{j-1} represents the cumulative amount of income declared up to the previous income class;
  - $\sum_{j=1}^{k} a_j$ is the total amount of income declared within a given municipality;
  - N is the total number of taxpayers in the municipality.

[9]The economic data derived from the database on municipal incomes of the Italian Ministry of Economy and Finance (MEF). For the social, demographic, and environmental variables, data are sourced from the Italian National Statistics Office (ISTAT). The characteristics of local administrators are elaborated from the dataset of the Ministry of the Interior. For Citizenship Income we use data from the Italian National Institute of Social Security (INPS).

[10]Notice that, according to ISTAT (2024) the share of people with an income of less than 10,000 euros and the average per capita income based on income declarations are officially classified as economic wellbeing indicators and are used to build the national informative system for planning, management and coordination of local authorities.

- *Average per capita income*: computed as the ratio between the total amount of income declared in each municipality and the total number of taxpayers. The formula used is:

$$I_{avg} = \frac{\sum_{i=1}^{n} A_i}{n} \tag{3.2}$$

where:
  - A_i is the income declared by individual i;
  - n is the total number of taxpayers in the municipality.

- *Share of poverty*: share of the population with an income of fewer than 10,000 euros (i.e., the ratio between the number of individuals who declare an income between 0 and 1000 thousand euros and the total of the declaring population). The formula is:

$$P_{share} = \frac{n_{\text{low income}}}{n_{\text{total}}} \tag{3.3}$$

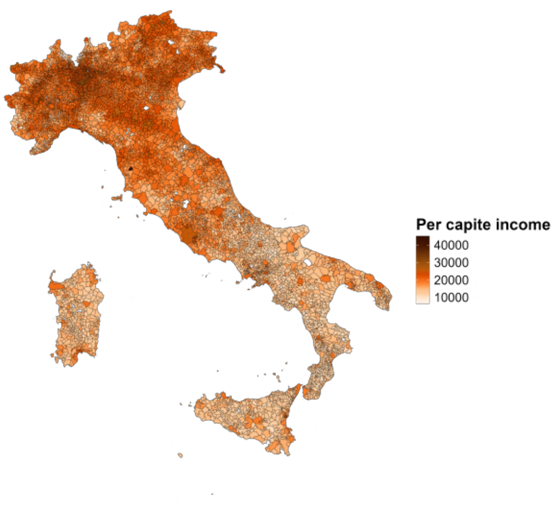
where:
  - $n_{\text{low income}}$ is the number of individuals with income less than 10k euros;
  - n_{total} is the total number of individuals who declare an income between 0 and 10k euros.

A municipal map for each indicator is illustrated in Figure 3.2.

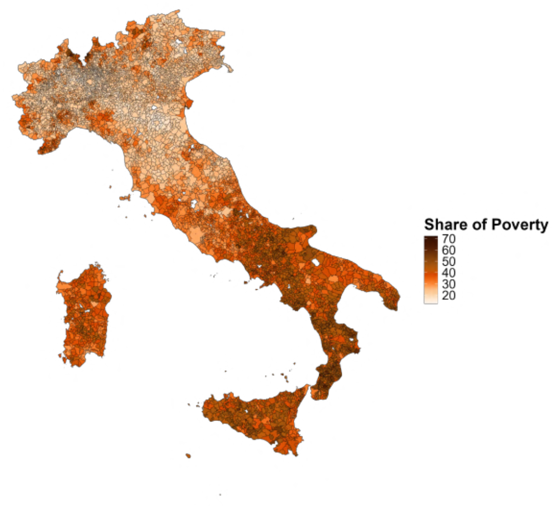
In our application, we analyse the time frame 2018-2022, that is, starting from the year preceding the approval of the GMI. The panel also includes 2018 because the three socio-economic variables will be lagged by one period with respect to the years of implementation of the GMI. Other controls will consider all the years relating to the operation of the measure (i.e., from 2019 to 2022).

Figure 3.2: Municipal Spatial Distributions of Economic Indicators (Average 2018-2021)

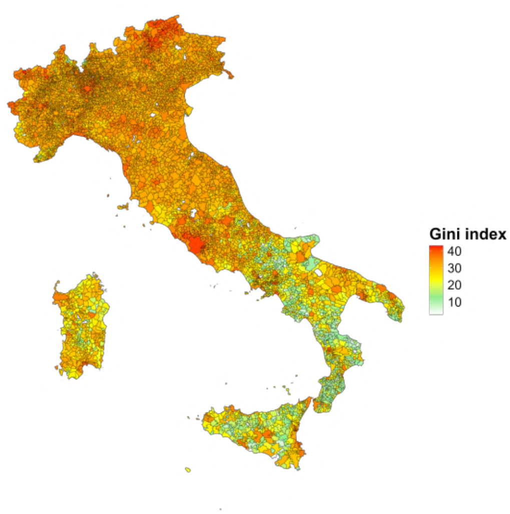
(a) Per Capita Income



(b) Poverty Share



(c) Gini Index



Source: Author's processing of MEF data.

3.3.2 Descritive Statistics

Table 3.1 reports the municipal descriptive statistics on these variables, aggregated on a macro-regional territorial scale, NUTS-1 level (Eurostat, 2023). This subdivision allows us to observe, at the large-scale, the clear spatial heterogeneity characterising the GMI phenomenon during the operation period, providing a reliable comparison between the areas of the country.

In terms of numbers of households receiving GMI we can detect a high average municipal difference between the North and the South. On average, in the municipalities of the North-East, approximately seventy families have benefited from the GMI in the years of implementation of the policy. In contrast, in the southern and island regions, much higher values are recorded, above the three hundred units at the average municipal level. Opposite trends can be seen from the observation of economic data, since greater average municipal wealth is recorded in the more advanced areas, where there is also greater variability in incomes (and therefore greater within disparity). Socioeconomic disparities, on a macro-regional scale, can be found in access to essential services and in data on the labour market, with the employment rate among those of working age being very marked between the North and the South (around 20 points). Although less spatially heterogeneous in the North-South dichotomy, inequalities are also present in terms of production plants at local level, with the Italian municipalities of the Central regions recording the highest number of production sites and employees at municipal level (300 units more local than in the North-East, 500 compared to the municipalities of the South and 600 compared to the North-West), as they are territories characterised by successful experiences in specific sectors of manufacturing and craftsmanship. The differences between the share of the working-age population with a level of elementary education are very marked, with much lower average values in the municipalities of the Southern regions and on the Islands. At a demographic level, the population growth rate demonstrates the long-term depopulation that afflicts the South, showing very negative municipal average values in these regions. In the North, and also in the municipalities of Central Italy, there is, instead, a high growth in the population. In the North-East the average municipal figure is close to 10%. The share of immigrants also follows similar trends, with the distribution of foreign citizens being concentrated mainly in the Centre-North. We also considered time-invariant variables describing the geographical characteristics or the elementary indicators of fragility at the local level. For instance, the Composite Index of Municipal Fragility (IFC) consists of 12 elementary indicators of a socioeconomic and environmental nature and represents the level of fragility of the municipalities and is used to study territorial risk factors. The geographical and environmental variables highlight values that vary territorially, depending on the nature of the indicator considered, with the Southern municipalities which, on average, present greater morphological problems, but fewer environmental problems, with less production of undifferentiated waste, and a lower incidence of hydro-geological risk, also connected to lower land consumption. Finally, there are no significant differences between the territorial areas in terms of characteristics of municipal administrators, such as age of the mayor and average age of councillors, suggesting a certain stability in local governance.

In general, these data reflect the economic disparities present in Italy, underlining the importance of targeted public policies and specific interventions, in order to reduce inequalities, particularly in the Southern and Island regions, where socioeconomic critical issues require urgent attention. At the same time, they highlight the need for an accurate evaluation of the protection programmes implemented.

Table 3.1: Descriptive Statistics on Socio-Economic Variables

	Years	North-East	North-West	Center	South	Islands
Economic						
Procapite Income	2018-2021	19834.44(2430.76)	19893.38(3126.37)	17575.90(2337.60)	14108.64(2151.76)	14062.89(2069.21)
Share of Poverty	2018-2021	24.55(4.791)	24.67(5.941)	29.81(5.176)	42.04(6.674)	40.84(5.892)
Number of Taxpayers	2018-2021	6390.86(16645.69)	3933.43(24256.14)	8696.04(63885.52)	4730.59(15285.53)	5120.31(16903.31)
Social						
Number of households	2019-2022	71.41(319.59)	73.79(875.02)	221.59(2192.98)	311.27(1931.18)	402.68(2117.39)
Local Units	2018-2021	729.53(2264.56)	469.51(4321.81)	1072.43(9261.02)	528.28(2200.39)	535.33(2054.45)
Employees of Local Units	2018-2021	2997.74(9139.23)	1880.05(18797.93)	3730.31(34266.97)	1589.83(7452.67)	1521.40(6753.87)
Employment rate 20-64 years	2019	72.94(3.982)	70.34(5.211)	65.24(6.048)	53.89(5.834)	51.96(5.637)
Distance to Essential Services	2019	29.20(18.35)	25.78(14.51)	30.85(15.61)	36.78(18.80)	42.29(23.73)
Demographic						
Resident population	2019-2022	8383.46(21859.15)	5331.30(32777.14)	12216.63(91763.44)	7623.43(27184.59)	8416.15(29656.03)
Share of Immigration	2019-2022	7.92(3.893)	7.717(4.352)	8.74(3.855)	4.366(3.012)	2.85(2.605)
Register for Foreigners	2019-2022	33.21(89.04)	19.68(122.36)	38.45(210.85)	20.26(59.22)	24.87(77.04)
Population Growth Rate	2019-2022	9.587(42.55)	6.359(52.28)	-2.845(42.94)	-28.01(48.35)	-25.46(42.01)
Population Dependency Index	2019-2022	72.13(9.156)	73.97(10.35)	73.20(8.726)	71.11(10.72)	71.45(8.58)
Population Density	2019-2022	1.87(3.522)	2.58(7.685)	2.234(3.359)	2.06(2.812)	2.506(2.889)
Population 25-64 with Low Shool	2019	35.13(6.529)	39.40(7.31)	37.04(6.304)	41.33(7.846)	50.12(8.312)
Geographic						
Level of Urbanisation	2019-2022	2.60(0.5196)	2.56(0.566)	2.682(0.4959)	2.60(0.5892)	2.664(0.5016)
Territorial Surface	2019-2022	44.73(47.04)	19.35(22.58)	59.94(73.62)	41.30(49.93)	65.17(72.01)
Altitude of the Center	2019-2022	305.24(368.43)	346.85(289.60)	355.56(222.74)	411.59(284.07)	336.35(259.51)
Environmental						
Protected Natural Areas	2019	14.27(22.30)	11.22(21.11)	18.79(25.33)	24.88(31.79)	17.19(24.26)
Landslide Hazard Zones	2019	6.187(11.24)	7.463(15.08)	9.76(10.74)	12.33(16.63)	4.813(7.746)
Land Consumption	2019	10.56(8.218)	11.87(11.49)	7.253(6.323)	9.19(11.02)	6.514(7.583)
High Emission Motor Rate	2019	18.95(4.908)	19.95(6.25)	26.37(7.206)	31.77(7.522)	33.46(8.131)
Undifferentiated Urban Waste	2019	135.69(105.91)	172.89(128.89)	191.01(138.98)	163.20(104.56)	132.88(95.08)
Administrative						
Age of Mayor	2019-2022	56.51(9.122)	58.64(10.44)	56.95(9.791)	56.79(8.977)	56.31(9.516)
Average Age of Councilors	2019-2022	51.36(4.29)	52.96(5.19)	51.67(4.459)	50.34(4.392)	48.87(4.411)

Source: Author's calculations based on INPS, Ministry of Economy, ISTAT and Ministry of the Interior data
Note: The values for the 2022 year relating to the "Number of Taxpayers", "Local Units" and "Employees of Local Units" are obtained as the average of the previous years due to the lack of available data. The Gini index is not present because it was calculated only at municipal level for econometric estimates.

3.4 Estimating Spatially-Varying Relationship via Spatially Clustered Regression

In the motivation Section (3.1.1), we stated that our goal was to investigate the spatial distribution of the participation rate of families involved in the GMI across the Italian municipalities and its relationship with a set of socioeconomic variables related to local wellbeing indicators, namely the Gini index, average per capita income, and the share of the population living in poverty (ISTAT, 2024). Motivated by the strong geographical heterogeneity of Italian macroeconomic and socioeconomic data, we might expect that this relationship is not uniform throughout the country, but exhibits complex spatial and temporal patterns. Therefore, we employ a spatial regression technique denoted as Spatially Clustered Regression (henceforth, SCR) that allows us to estimate spatially-varying empirical relationships in which the coefficients linking the response variable and the covariates are grouped into internally homogeneous spatial clusters. This technique was introduced by Sugasawa and Murakami (2021) as an alternative method to the Geographically Weighted Regression (GWR) of (Brunsdon et al., 1998; Fotheringham et al., 2022). Both techniques are suitable tools to address the spatial heterogeneity issue (Zhu and Turner, 2022), that is, empirical frameworks in which the relationship between variables is not constant over space. While GWR, and its variants, such as Geographically Weighted Regression Lasso (Wheeler, 2009) and Geographically and Temporally Weighted Regression (Wu et al., 2014), frequently lead to unstable estimates that are sensitive to the size of the dataset being considered. SCR provides a good compromise between computational efficiency, stability of results, and interpretability. Moreover, SCR naturally adapts to the case of Generalized Linear Models (GLMs), see Section 15 of Fox (2015), enabling the modelling of more complex data structures than is possible with the Gaussian case and also has the potential to be expanded to more fine-grained modelling, such as Generalized Additive Models (Wood, 2020).

SCR combines linear regression models with spatial clustering; for an extended review of spatial clustering algorithms (Kopczewska, 2022) of cross-sectional units based on the idea that the relationship between covariates and response variable for nearby observations is similar (or even identical), but that this relationship might vary between spatially distant groups of observations. In practice, SCR assumes that (1) units can be divided into a finite number of spatial clusters, where units in the same groups share the same regression coefficients; (2) group membership is based on the idea that nearby or neighbouring geographic units are likely to belong to the same groups (Potts, 1952). Technically, clustering is performed through a penalised version of the iterative K-means algorithm in which spatial proximity between observations is taken into account to form groups, favouring the clustering of neighbouring units in space. See Wang et al. (2021) for a similar idea of creating non-overlapping spatial partitions of locations in a model validation context. As the SCR treats the spatial dependence using a spatial weighting matrix, such dependence can be either modelled through a spatial contiguity matrix or a distance matrix; see, for instance, Section 4 of Kopczewska (2020). A potential limitation of the SCR approach is that it allows only for cross-sectional (spatial) estimates and not spatio-temporal specifications. In the following Section 3.5, we provide an empirical strategy to overcome this drawback.

Let us denote the observed response variable at location s as y_s and denote the $(p \times 1)$ vector of covariates at location s as $\boldsymbol{x}_s$, where $s = 1, \ldots, n$ is the index for the locations and n is the overall number of locations. In the present case, the locations are the $n = 7850$ Italian municipalities.

Let us suppose that, conditioning on the p exogenous covariates, the response variable follows a certain conditional distribution:

$$f(y_s|\boldsymbol{x}_s; \boldsymbol{\theta}_s) \tag{3.4}$$

where:

- $\boldsymbol{\theta}_s$: is a vector of unknown parameters to be estimated.

Typically, the distribution $f(\cdot)$ is a member of an exponential family, such as the Gaussian, the Poisson or the Binomial distributions. In the GLMs framework, such an assumption implies that the expected value of the response variable (i.e., $E(y_s) = \mu_s$) and the p covariates are linearly related through a linear link function $\eta_s = g(\mu_s)$ transforming the expectation of the response variable into the linear predictor, that is:

$$\eta_s = \theta_{0s} + \boldsymbol{\theta}_s \boldsymbol{x}_s \quad \forall s = 1, \ldots, n \tag{3.5}$$

where:

- $\boldsymbol{\theta}_s$: is the vector of p unknown regression parameters;
- θ_{0s}: is the unknown intercept.

Due to the spatial heterogeneity, for each observation s, we might assume that each location s is associated with a location-specific set of regression parameters. However, the model would suffer from an identification problem, making its estimation unfeasible. Therefore, we assume that the n locations can be divided into a finite number $G < n$ of clusters and that observations within each group $g = 1, \ldots, G$ share the same parameter values. Therefore, Equation 30 reduces to:

$$\eta_{s_g} = \theta_{0g} + \boldsymbol{\theta}_g \boldsymbol{x}_{s_g} \quad \forall g = 1, \ldots, G \tag{3.6}$$

where:

- $\boldsymbol{\theta}_g$: is the vector of p cluster-specific regression parameters;
- θ_{0g} is the cluster-specific intercept;
- $\boldsymbol{x}_{s_g}$: represent the value of the covariates;
- y_{s_g}: represent, respectively, and response variable for the n_g locations belonging to the g-th cluster;

- $\eta_{s_g} = g(\mu_{s_g})$: is the link function for g-th cluster. Notice that we assume different group sizes and group-wise parameters, but a common link function.

GLMs are usually fit via maximum likelihood (ML), which provides both estimates of the regression coefficients and the corresponding asymptotic (i.e., large-sample) standard errors. The membership of a unit to a certain group is unknown and must be estimated. In order to have groups consisting of spatially close (or even contiguous) units, following the approach of Potts (1952), the likelihood function to be optimised is augmented by a penalty term that induces neighbouring units to clump together. Maximum likelihood estimation of the group-wise parameters and the membership is accomplished by employing a K-means-like iterative algorithm that iteratively updates the membership and the local regression parameters. At each iteration, the update consists of maximising the log-likelihood function based on the classified units in each cluster. Computational details of the algorithm are described in Section 2 (Algorithm 1) of Sugasawa and Murakami (2021). As a further remark, as the estimation is performed under a likelihood paradigm, model selection (i.e., to identify the optimal number of groups) can be performed using information criteria, such as the Akaike's Information Criterion (AIC) and the Bayesian Information Criterion (BIC) (Sugasawa and Murakami, 2021; Di Mari et al., 2023).

3.5 Empirical Strategy

3.5.1 Model Specification

Following the methodological design illustrated in Section 3.4, we adopt an empirical strategy that considers as the response variable the number of households benefiting from the GMI at the municipal level for Italy. Figure 3.1 shows a dynamic municipal mapping of the proportion of families which benefited from the GMI intervention out of the total number of families in each municipality. The number of families is actually a count variable that takes integer values only. Available municipal data for the GMI policy span the period 2019 to 2022.

The natural choice of GLM for count data is the Poisson distribution. Specifically, we adopt a Poisson regression approach with logarithmic link function, which has, as its main advantage, the capability of interpreting the estimated regression coefficients as a percentage change in the response following a unit increase in the covariate. Given the generic covariate x_{js_g} for group g, with $j = 1, \dots, p$ and $g = 1, \dots, G$, and the corresponding coefficient estimate $\hat{\beta}_{jg}$, the value $100 \times \hat{\beta}_{jg}$ represents the estimated local (i.e., for the g-th spatial cluster) percentage variation in the number of GMI-receiving households associated with a unit change in x_{js_g}.

As previously stated, we use a large set of municipal-level characteristics from 2018 to

2022 where the main drivers of the analysis are three variables related to local wellbeing, namely the municipal Gini index, the average municipal per capita income and the poverty share (share of the population with an income below 10,000 Euros). To obtain a relevant interpretation of the GMI participation rates, these three indicators are included in the model with a lag of one period relative to the years of operation of the policy (e.g., one using the GMI for 2021, the per capita income variable refers to the year 2020). This choice is motivated by the structural characteristics connected to the socio-economic criticality, which impact, in a lagged and non-simultaneous fashion, the requests for income support. In addition, by construction, the policy refers to an income condition (ISEE) delayed by one year concerning the request for subsidy. Furthermore, to avoid spurious results, we introduce into the regression model a set of socio-demographic, geographical, environmental, and institutional control variables, temporally aligned with the years of activation of the policy. Overall, $q = 21$ control variables were included in the model. These variables are reported in Table 3.1.

Even though it is not modelled explicitly, we take into account the temporal dynamics of the relationships by considering more than one year of municipal data. Specifically, SCR is independently applied to data for the years 2019, 2020, 2021, and 2022[11]. The temporal evolution of the estimated local coefficients, municipal membership and number of groups will provide insights into how and whether the phenomenon of income support claims has changed or settled.

The empirical specification of the Poisson regression model for the generic cluster g, composed of the units s_g, and the generic year t is described in Eq. 32:

$$\begin{aligned} log[E(y_{s_g t}|\boldsymbol{X}_{s_g t})] &= \beta_{0gt} + \beta_{1gt} Gini_{s_g t-1} + \beta_{2gt} PerCapitaIncome_{s_g t-1} + \\ &= \beta_{3gt} ShareOfPoverty_{s_g t-1} + \boldsymbol{\gamma}_{gt} X_{s_g t} \\ &\quad \forall t = 2019, \ldots, 2022 \text{ and } \forall g = 1, \ldots, G \end{aligned} \tag{3.7}$$

where:

- $y_{s_g t} \sim Poisson(\eta_g)$: are the observed number of recipient families for year t in group g;
- β_{1gt}: is the coefficient associated with the income inequality (Gini index) for group g at year $t-1$;
- β_{2gt}: is the coefficient associated with the per capita income for group g at year $t-1$;
- β_{3gt}: is the coefficient associated with the share of poverty for group g at year $t-1$;
- $\boldsymbol{\gamma}_{gt}$: is the set of $q = 21$ contemporaneous coefficients associated with the cluster-specific control variables.

In the following, we will refer to Equation 32 as *Specification 1*.

[11] Notice that, due to abolitions and fusions among municipalities, the yearly number of units varies across the period. In the present study, the number of municipalities for 2019 is 7862, for 2020 is 7829, for 2021 is 7848, and for 2022 is 7895.

The empirical strategy adopted has several merits. First, it allows us to identify the actual effect of global and local socioeconomic weaknesses on citizens' participation in cash benefits for income support, taking into account other exogenous factors. Second, it allows us to explicitly take into account the spatial heterogeneity and, therefore, to study any territorial and regional dependencies, in relation to the number of households benefiting from the measure. Third, we are allowed to study the temporal dynamic of the estimated relationships by comparing the four separate models

3.5.2 Robustness Checks

As a robustness analysis, we repeat the estimates with respect to different specifications of the main model in 32. For each robustness specification, we consider a subset of three lagged socioeconomic variables listed above (i.e., we estimate models with only one of three or with pairs of variables) accompanied by the relevant controls. In this way, we obtain a direct comparison useful for determining the weighting that each specific income variable has on obtaining the GMI. Overall, five empirical specifications are estimated, of which *Specification 1* (i.e., Eq. 32) is used as a benchmark model. For the sake of brevity, the equations of Specifications 2 to 5 are reported in the Appendix 3.7.

A further issue to be considered is the way in which spatial dependence is treated. The spatial dependence in an SCR model is modelled through a spatial contiguity matrix or a distance matrix. Since we are dealing with areal data (lattice), in our case, we prefer to use a contiguity matrix in which the number K nearest neighbours is fixed. Considering the high number of cross-sectional units at our disposal, we test three different specifications of the contiguity matrix: $K = 25$, $K = 50$ or $K = 100$ neighbors.

3.5.3 Model Comparison and Selection

The identification of the best model for each year and specification is based on statistical goodness-of-fit measures. To this purpose, we use the Bayesian Information Criterion (BIC), that, when minimised, provides a statistical insight into the optimal number of groups/clusters and the corresponding estimate of the local parameters (Di Mari et al., 2023). In the context of SCR, in fact, Italian municipalities are categorised into internally homogeneous clusters, ensuring that the influence of economic indicators remains consistent within each group, while varying between different spatial groups.

The choice of the BIC compared to other criteria is motivated by its parsimony property, which, in this context, leads to preferring a smaller number of groups, improving the overall interpretability of the socioeconomic phenomena under investigation. To this extent, consider the following: In Italy there are approximately 8 thousand municipalities. When considering a classification with $G = 25$ groups, there would be on average 320 municipalities in each cluster; becoming 160 when the number of groups is increased to $G = 50$ With the intention of preserving the asymptotic statistical properties of the regression estimators, it is necessary to maintain the highest possible sample size in each cluster. Therefore, for each specification, year, and number of neighbours, we estimate the five specifications with a number of groups ranging from $G = 1$ to $G = 50$, where $G = 1$ indicates the pooled or global regression (therefore, without local effects). However, it is worth noting that as the number of groups increases, the understanding of the estimates reduces. In this sense, BIC minimisation guarantees an adequate compromise between complexity and interpretability.

Overall, the total number of econometric models estimated is equal to 3000 (given by 4 years $\times$ 50 clusters $\times$ 3 neighbouring structures $\times$ 5 specifications). Despite the computational burden and the use of automatic selection criteria, the visual inspection of the results and their economic interpretation act as the main guidelines for the evaluation of the results.

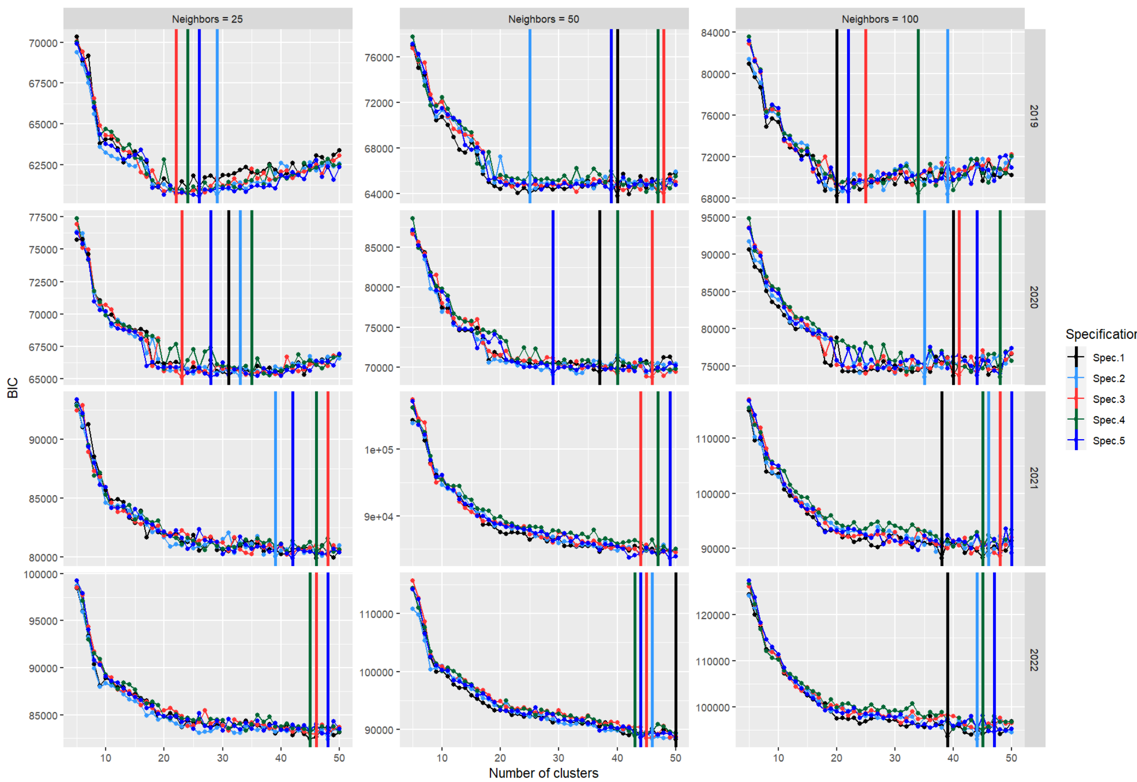
3.6 Results

3.6.1 Optimal Number of Groups and Complexity of the Social Phenomenon

Figure 3.3 shows, for every year of GMI implementation, that the optimal model is associated with a number of groups between 25 and 48. The BIC, at least for 2019, shows a parsimonious approach for establishing the optimal number of groups (25 groups with about 320 municipalities in each group), as adding additional groups would not significantly improve the model fit. The observation of the BIC highlights an additional aspect of considerable relevance, pertaining to the optimal group number. One can notice that, moving from one year to the next, the number of optimal groups almost doubles from 25 (2019) to 48 (2022). Such increases indicate the growth in the complexity of the underlying phenomena. The increase in the degree of complexity is directly linked to the increase in spatial heterogeneity in the investigated relationships. In the following subsections 3.6.2, 3.6.3, 3.6.4, we analyse whether this increase is associated with a change in the signs and magnitudes of the estimated relations.

At a general level, empirical results obtained in our benchmark specification (Eq. 32) reveal the historical geographical dualism characterising the Italian socio-economic development, that is, the one between Northern and Southern regions. At the same time, they also show marked differences within the same territories and regions. In particular, they highlight a positive correlation between low levels of municipal per capita income and high local share of poverty concerning the number of beneficiary families. Spatially discordant values, due to the structural characteristics of the Italian economy are, instead, observed in areas characterised by high income inequalities. In other words, a greater territorial inequality can influence the participation rates of families and individuals in income support policies since, in these areas, the average level of income is higher. Similar results are also obtained in the other specifications of the model, formally illustrated in the equations of Appendix 3.7.

Figure 3.3: Bayesian Information Criterion (BIC) by Number of Clusters, Neighbors, and Year



3.6.2 Heterogeneity in Income Inequality

Estimates of the coefficients associated with income inequality, embedded in the municipal Gini index, show non-uniform spatial influence, especially in the Northern regions (see Figures 3.4a and 3.4b). Indeed, we obtain negative and partly significant correlation between income inequality and GMI attainment in the North-West and North-Central regions. In contrast, some areas in the North-East, including Veneto and Friuli-Venezia Giulia, appear to be an interesting case for assessing the relationship between income inequality and monetary benefit, as a strong positive connection emerges between income inequality and beneficiary households. Indeed, these territories have the most positive and significant coefficients in Italy. In the Central and Southern Regions, i.e., the municipalities with the lowest values of income inequality, on the other hand, the results show a positive correlation between inequality and recipient households.

The differences in the estimates found among the northern regions are very important for assessing the influence of inequality on benefit receipt, since the greatest economic inequality is concentrated in these areas (see Figure 3.2c showing Gini Index values at the municipal level). These differences, both in sign and direction, could be attributed to several factors, such as the heterogeneous degrees of the quality of institutions or the presence of social protection and social assistance systems on a local or regional scale at (Nifo and Vecchione, 2014; Rodríguez-Pose, 2013; Gallo, 2021). More generally, they may reflect the structural conditions of the Italian economy, which can assign different influences to income inequalities depending on the territorial specificities in which they develop.

As mentioned earlier, it is worth noting that the estimates change significantly between 2019 and 2022. In particular, the degree of complexity of the relationship increases, as well as the spatial heterogeneity. The increase in complexity is also associated with a change in the signs of the coefficients of income inequality, which, in general, move from negative to positive, suggesting a greater direct correlation between income inequality and the receipt of the GMI monetary transfer. For example, considering the estimates associated with the Gini index in 2021 or 2022, thirteen clusters show a positive and statistically significant coefficient (while one six groups have a negative and significant value). In 2019, the number of clusters with positive and significant coefficients were eight. This fact is consistent with the hypothesis of an increase in the strength of the relationship between income inequality and number of benefiting families. One potential motivation for this result would relate to the recessive effects of the COVID-19 pandemic, which has widened the inequalities that were already existing in the country (Palomino et al., 2020; INPS, 2022a,b; Gallo and Raitano, 2023).

3.6.3 Heterogeneity in Per Capita Income

Figure 3.5a shows the coefficients of *Specification 1* relative to municipal per capita income levels. In the Northern regions, negative and statistically significant coefficients are found throughout the observed series, particularly in the municipalities of the North-East. These municipalities, which are characterised by the highest levels of municipal income per capita (as highlighted in Figure 3.2a and Table C2), show the most negative correlations between average income level and number of GMI recipients. Negative and significant coefficients are also present in the middle and upper Tyrrenian regions (Lazio and Tuscany). Positive correlations, on the other hand, are observed among the municipalities of the Adriatic regions and in the territories of the South and Islands, where the lowest levels of per capita income are present. Therefore, the results show that the different levels of per capita income found in the Italian territory are properly correlated with the number of households benefiting from the IMG monetary transfer.

The results suggest an effective targeting of the policy in areas of the country with lower income levels. Referring to the above-mentioned increase in the degree of complexity and spatial heterogeneity of socioeconomic vulnerability, we can also detect an increase in the correlation between the number of recipients and social exclusion situations during the reference period, evidenced by an increasing number of positively estimated coefficients with higher and higher magnitudes. In particular, the maps in Figure 3.5a and 3.5b show that, for 2019, an almost exact overlap exists between the spatial boundaries created by the clusters and the administrative boundaries of the Italian regions. In 2022, however, the increase reveals deeper spatial differences. Again, the weight of the pandemic may have been decisive in the link between income differentials and the number of households benefiting from the measure (Palomino et al., 2020; Brunori et al., 2021; INPS, 2022a,b; Gallo and Raitano, 2023).

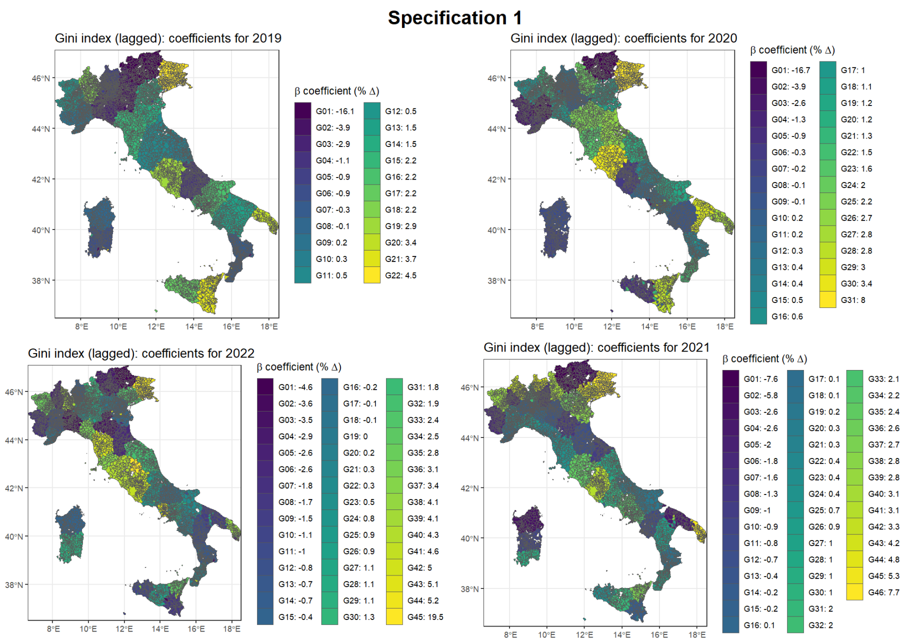
3.6.4 Heterogeneity in the Share of Poverty

Estimates of the coefficients associated with the proportion of municipal people living in poverty (see Figures 3.6a and 3.6b) show results consistent with the Italian spatial dynamics (shown in Figure 3.2b and Table C2). Indeed, in the municipalities of Northern Italy, there is a largely negative and significant correlation between citizens in poverty and households receiving Citizenship Income. In particular, in the municipalities of the North-East, the values of the coefficients remain fairly constant in terms of sign, direction, and intensity, over the four years of observation, showing, in this sense, some stability. Similar results, but with opposite signs, are recorded in the municipalities of the South and Islands, where the coefficients of the Eq. 32 show, instead, a positive and statistically significant link between poverty and receipt of the monetary benefit.

The results obtained on share of poverty are thus in line with the well-known socioeconomic vulnerabilities present in the Italian territory. Consequently, the share of poverty conditions seems to be well related to the number of households benefiting from the GMI. Even in the case of the municipal poverty share, the growth of complexity along the observed annuals is revealed by the occurrence of coefficients having more positive magnitudes, especially in Southern Italian municipalities. Specifically, in the last year considered, numerous socio-spatial differences related to social exclusion phenomena can be observed even in similar geographical contexts.

Figure 3.4: Regression Analysis: Beta Coefficients of Gini Index by Year of Operation of the GMI

(a) *beta* Coefficients



(b) t-test

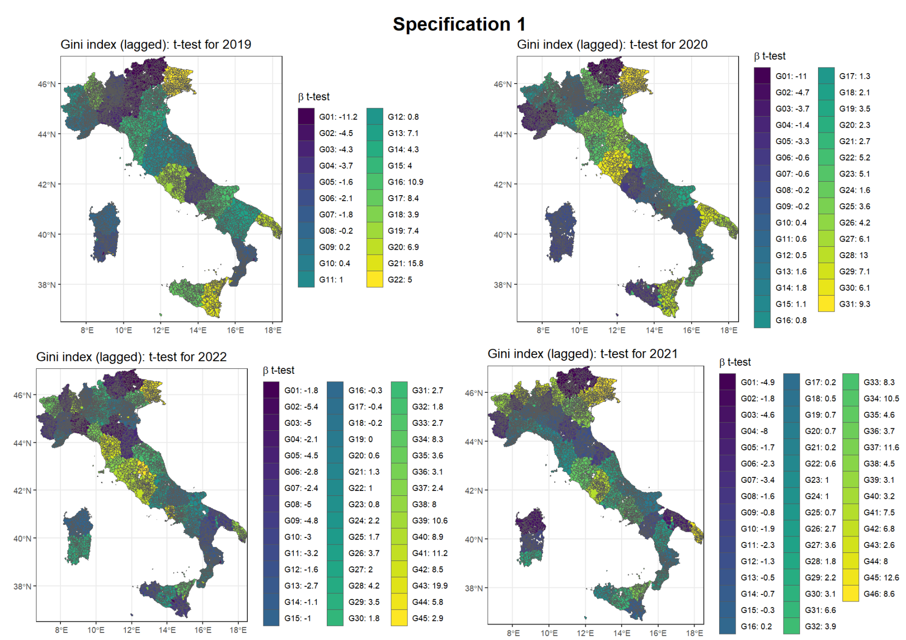
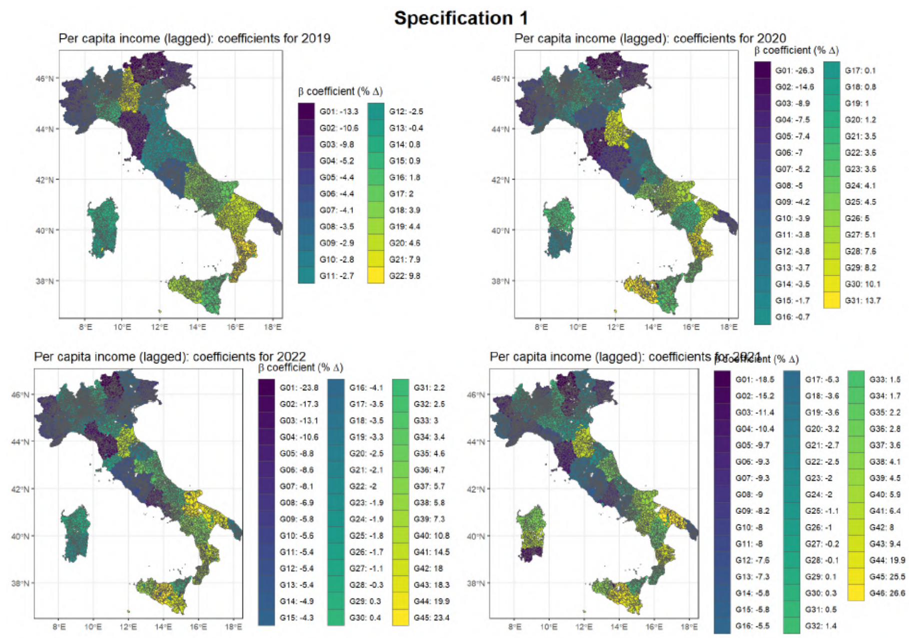


Figure 3.5: Regression Analysis: Beta Coefficients of Per Capita Income by Year of Operation of the GMI

(a) *beta* Coefficients



(b) t-test

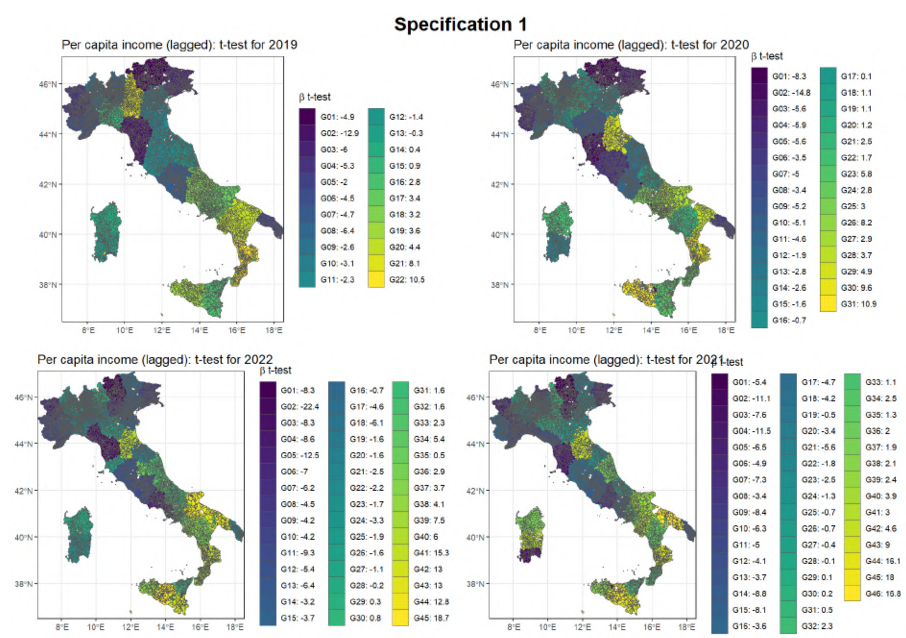
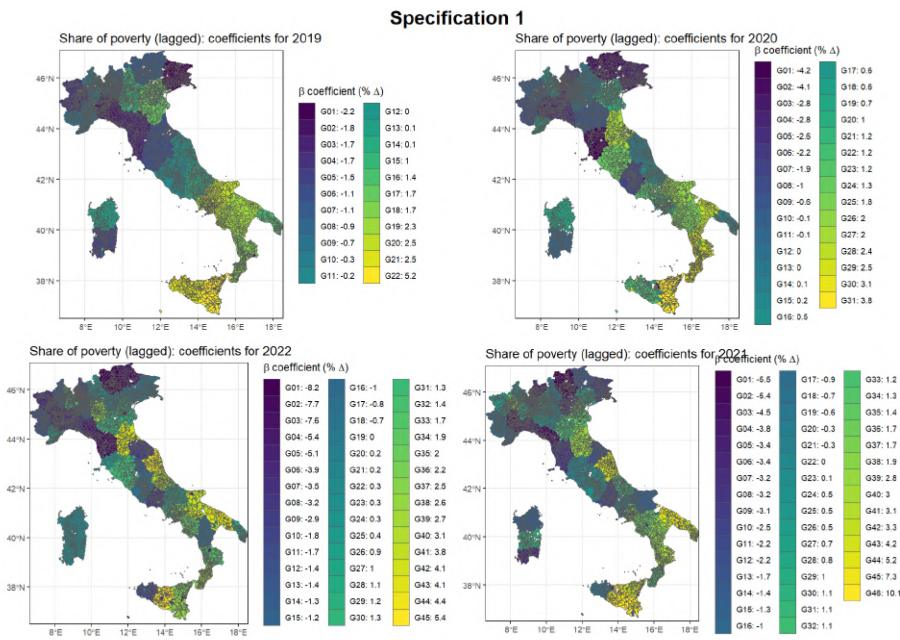
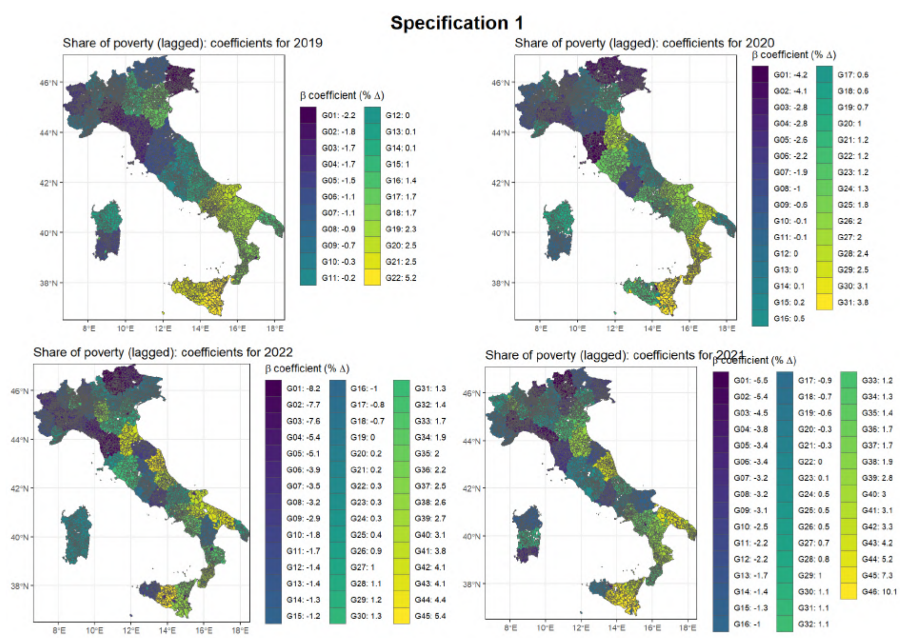


Figure 3.6: Regression Analysis: Beta Coefficients of the Share of Poverty by Year of Operation of the GMI

(a) *beta* Coefficients



(b) t-test



3.7 Concluding Remarks

In this paper, we investigated, at the municipal level, the role of three local socioeconomic and wellbeing indicators, that is, the Gini index for income inequality, the average per capita income, and the share poverty (number of citizens with an income below 10 thousand €), in determining geographical differences and heterogeneity in the number of families that participate in income support from central government.

In particular, we considered a guaranteed minimum income (GMI) support policy implemented in Italy to support families and individuals affected by socioeconomic exclusions, namely the Italian Citizenship Income (in Italian language, *Reddito di Cittadinanza*) adopted between 2019 and 2023. The main research question focused on whether the GMI policy was able to effectively reach the targeted families in light of the geographical patterns shown by the three local indicators.

To answer the question we implemented numerous econometric specifications with spatially-varying coefficients, namely the spatially-clustered regression (SCR) models, grasping the strong spatial heterogeneity exhibited by the GMI recipient households by grouping municipal units into homogeneous and spatially-contiguous groups. In this way, we were able to evaluate how geography and local factors influence the effective coverage of income support policies. Within the SCR approach, the response variable was the count of households benefiting from the policy, while the local socioeconomic and wellbeing indicators acted as exogenous variables, in addition to a large set of socio-political and economic controls. To capture the lagged but structural effects of these socioeconomic indicators, the variables were lagged by one year relative to policy implementation in the models.

Our findings showed that the spatial heterogeneity of socio-economic proxies across the Italian municipalities significantly influenced the participation for income support. Indeed, both the sign and magnitude of the estimated correlation strongly depend on the type of indicator used and by territorial (i.e., local) structural characteristics, as well as by the year under inspection. We obtained positive and statistically significant correlations regarding the level of per capita income and the share of municipal poverty. In particular, the highest positive magnitudes were recorded in areas marked by higher socioeconomic issues and low-income levels (i.e., Southern Italy). Concerning income inequality, we obtained spatially-varying but representative results. Specifically, we assessed the simultaneous presence of areas characterised by marked income inequalities which present negative correlations with the number of beneficiary families (e.g., North-Western Italy), and areas showing positive and statistically significant values of correlation (e.g., North-Eastern Italy). These results, which may seem contradictory, are intrinsically linked to the examined dimensions. In fact, in areas where both average per capita income and income inequality are high (e.g., North-West), the GMI, which by definition applies only to low-income households, was unable to reach potential household targets, leaving the level of income inequality unchanged and leading to a negative correlation. In contrast, in Southern Italy, mainly characterised by low income levels but lower levels of income inequality, the GMI reached a high number of households, leading to a positive and significant correlation. Eventually, the results highlighted a remarkable augmentation in the complexity of the social phenomenon throughout the period 2019-2022. Such a complexity was proxied by the number of relevant groups identified through the SCR algorithm. Empirically, we noticed that, from the second year of activity of the GMI policy (i.e., 2020) onwards, with the arrival of the COVID-19 pandemic, the number of clusters with statistically significant but spatially-varying coefficients increased substantially, leading to a more scattered and complex scenario.

The Italian Citizenship Income has been charged with numerous criticisms, mainly focused on the undefined effectiveness of the policy to properly identify and support families in need of assistance. This paper provides fresh and positive evidence regarding the primary goals of the GMI policy, that is, reaching and being helpful for families in socioeconomic exclusion, while accounting for territorial-specific socioeconomic features. The main conclusion we arrive at is that local socioeconomic conditions deeply matter when implementing national policies. Today, this measure has been abolished and replaced with other forms of income support that are less generous in terms of the amounts payable and the number of eligible families. In light of these regulatory changes, we sustain the need for the constant evaluation of income support policies, by exploiting the role of spatial heterogeneity to meet the non-homogeneous economic and social contexts typical of the Italian panorama.

Appendix C

C1. Descriptive Statistics on GMI and Socio-Economic Incators

Table C1: Distribution of GMI by Regions (2019-2022)

Territorial	**Number of households receiving GMI**				**Share households receiving GMI on total households**			
	2019	**2020**	**2021**	**2022**	**2019**	**2020**	**2021**	**2022**
Abruzzo	18197	23522	40448	43689	2.92	3.69	6.07	6.61
Basilicata	9176	10657	17083	19758	4.01	4.59	7.27	8.35
Calabria	58774	78214	135680	155902	6.69	8.78	14.60	16.92
Campania	160853	249541	450921	502973	5.43	7.82	13.22	14.56
EM	27252	36818	68091	66248	1.13	1.49	2.80	2.75
FVG	8979	11236	18432	18283	1.09	1.31	2.31	2.31
Lazio	73480	110533	221152	236378	3.53	4.92	8.57	9.21
Liguria	17129	24285	43747	42262	1.87	2.56	4.69	4.84
Lombardy	65013	94639	184317	168206	1.16	1.55	2.92	2.85
Marche	11974	15682	26923	26156	1.66	2.10	3.60	3.53
Molise	5077	6631	11238	12031	3.63	4.99	8.29	8.82
Piedmont	46069	64045	116903	117995	1.47	2.05	3.79	3.95
Puglia	78260	106819	194233	213635	4.12	5.73	9.87	10.65
Sardinia	37892	47248	79150	87146	5.06	6.11	9.61	10.08
Sicily	148898	215302	380350	438769	6.06	8.39	13.69	15.72
Tuscany	30955	39567	70199	69260	1.70	2.17	3.79	3.71
TAA	2405	3489	7584	6705	0.38	0.51	1.06	0.94
Umbria	8640	11748	19977	20515	1.92	2.58	4.42	4.51
Valle d'Aosta	908	1068	1812	1634	1.10	1.26	2.08	1.87
Veneto	23789	30818	55895	53232	0.94	1.19	2.09	2.04
North-East	62425	82361	150002	144468	0.89	1.14	2.08	2.02
North-West	129119	184037	346779	330097	1.34	1.82	3.38	3.42
Center	125049	177530	338251	352309	2.43	3.27	5.68	5.89
South	330337	475384	849603	947988	4.85	6.58	11.02	12.27
Islands	186790	262550	459500	525915	5.57	7.27	11.69	12.95
Italy	**833720**	**1181862**	**2144135**	**2300777**	**2.60**	**3.47**	**5.95**	**6.40**

Source: Author's processing of INPS data.
Note: The data are obtained by aggregating municipal data provided by the INPS.

Table C2: Average Per Capita Income and Average Share of Poverty by Regions (2018-2021)

Territorial	**Average Per capita income**				**Average Share of Poverty**			
	2018	**2019**	**2020**	**2021**	**2018**	**2019**	**2020**	**2021**
Abruzzo	14901.66	15057.54	15098.77	15760.44	37.86	37.18	37.18	35.28
Basilicata	13910.99	13965.20	14072.04	14720.24	41.51	40.70	39.90	38.08
Calabria	12730.35	12849.30	12867.45	13384.31	48.10	47.22	47.00	45.43
Campania	14198.66	14293.76	14218.47	14899.26	42.98	42.15	42.37	40.22
EM	19981.44	20161.18	19788.60	20872.62	23.51	23.10	23.43	22.31
FVG	19129.32	19313.80	19120.46	20008.52	25.43	24.97	25.08	24.01
Lazio	16686.23	16769.11	16701.81	17397.22	34.13	33.43	33.68	31.75
Liguria	18167.64	18143.74	17784.65	18729.52	29.40	29.02	30.12	28.52
Lombardy	20531.92	20535.43	20149.30	21225.59	23.70	23.46	24.36	22.94
Marche	16955.62	17173.78	17007.16	17937.82	29.07	28.33	28.39	26.87
Molise	13056.98	13165.56	13206.27	13807.45	44.98	44.01	43.70	41.90
Piedmont	19186.02	19225.76	18929.11	19918.88	25.48	25.18	25.52	24.29
Puglia	14124.38	14231.69	14257.25	14913.95	41.88	41.14	41.19	39.24
Sardinia	14011.69	14169.18	14206.48	14733.04	40.14	39.09	39.23	37.52
Sicily	13634.56	13753.58	13704.98	14325.24	43.65	42.83	43.08	40.92
Tuscany	18637.47	18771.97	18364.43	19480.18	27.11	26.58	27.32	25.71
TAA	19983.33	20161.27	19815.55	20659.54	27.22	26.52	26.53	26.27
Umbria	17306.11	17498.55	17398.27	17741.37	29.08	28.67	29.02	28.95
Valle d'Aosta	19885.15	19965.56	19665.50	20032.98	24.17	23.84	24.86	25.41
Veneto	19406.38	19577.21	19238.95	20280.69	24.75	24.24	24.56	23.43
North-East	19616.80	19792.29	19469.04	20454.86	25.06	24.55	24.77	23.84
North-West	19799.65	19817.63	19470.86	20485.68	24.86	24.58	25.28	23.97
Center	17357.98	17496.37	17307.41	18144.92	30.50	29.87	30.22	28.64
South	13867.21	13977.84	13985.49	14606.02	43.15	42.34	42.27	40.40
Islands	13819.67	13957.31	13951.65	14525.35	41.93	41.00	41.19	39.25
Italy	**17541.57**	**17634.11**	**17445.99**	**18295.21**	**31.40**	**30.85**	**31.13**	**29.69**

Source: Author's processing of Ministry of the Economy data.

Note: The data are obtained by aggregating municipal data provided by the Ministry of Economy.

C2. Additional Specifications and Results of the Model

Below the additional empirical specifications of the Poisson regression model for a generic cluster g and year t is described in Eq. 33, 34, 35, 36.

***Specification 2*:**

$$\begin{aligned} log[E(y_{s_g t}|\boldsymbol{X}_{s_g t})] = \beta_{0gt} + \beta_{1gt} PerCapitaIncome_{s_g t-1} \\ + \beta_{2gt} ShareOfPoverty_{s_g t-1} + \boldsymbol{\gamma}_{gt} X_{s_g t} \\ \forall t = 2019, \ldots, 2022 \text{ and } \forall g = 1, \ldots, G \end{aligned} \tag{3.8}$$

where:

- $y_{s_g t} \sim Poisson(\eta_g)$: are the observed number of recipient families for year t in group g;
- β_{1gt}: is the coefficient associated with the per capita income for group g at year $t-1$;
- β_{2gt}: is the coefficient associated with the share of poverty for group g at year $t-1$;
- $\boldsymbol{\gamma}_{gt}$: is the set of $q = 21$ contemporaneous coefficients associated with the control variables.

***Specification 3*:**

$$\begin{aligned} log[E(y_{s_g t}|\boldsymbol{X}_{s_g t})] = \beta_{0gt} + \beta_{1gt} Gini_{s_g t-1} + \boldsymbol{\gamma}_{gt} X_{s_g t} \\ \forall t = 2019, \ldots, 2022 \text{ and } \forall g = 1, \ldots, G \end{aligned} \tag{3.9}$$

where:

- $y_{s_g t} \sim Poisson(\eta_g)$: are the observed number of recipient families for year t in group g;
- β_{1gt}: is the coefficient associated with the income inequality (Gini index) for group g at year $t-1$;
- β_{2gt}: is the coefficient associated with the share of poverty for group g at year $t-1$;
- $\boldsymbol{\gamma}_{gt}$: is the set of $q = 21$ contemporaneous coefficients associated with the control variables.

***Specification 4*:**

$$\begin{aligned} log[E(y_{s_g t}|\boldsymbol{X}_{s_g t})] = \beta_{0gt} + \beta_{1gt} PerCapitaIncome_{s_g t-1} + \boldsymbol{\gamma}_{gt} X_{s_g t} \\ \forall t = 2019, \ldots, 2022 \text{ and } \forall g = 1, \ldots, G \end{aligned} \tag{3.10}$$

where:

- $y_{s_g t} \sim Poisson(\eta_g)$: are the observed number of recipient families for year t in group g;
- β_{1gt}: is the coefficient associated with the income per capita income for group g at year $t-1$;
- β_{2gt}: is the coefficient associated with the share of poverty for group g at year $t-1$;
- $\boldsymbol{\gamma}_{gt}$: is the set of $q = 21$ contemporaneous coefficients associated with the control variables.

***Specification 5*:**

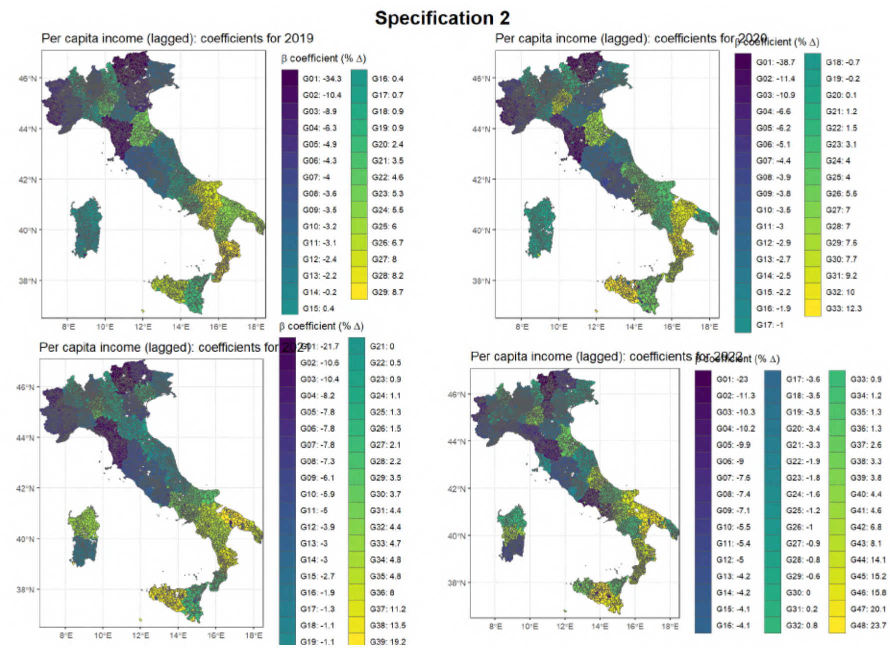
$$\begin{aligned} log[E(y_{s_g t}|\boldsymbol{X}_{s_g t})] = \beta_{0gt} + \beta_{1gt} PerCapitaIncome_{s_g t-1} + \boldsymbol{\gamma}_{gt} X_{s_g t} \\ \forall t = 2019, \ldots, 2022 \text{ and } \forall g = 1, \ldots, G \end{aligned} \tag{3.11}$$

where:

- $y_{s_g t} \sim Poisson(\eta_g)$: are the observed number of recipient families for year t in group g;
- β_{1gt}: is the coefficient associated with the share of poverty for group g at year $t-1$;
- β_{2gt}: is the coefficient associated with the share of poverty for group g at year $t-1$;
- γ_{gt}: is the set of $q = 21$ contemporaneous coefficients associated with the control variables.

Figure C1: Regression Analysis - *Specification 2*: Results of Per Capita Income by Year of Operation of the GMI

(a) *beta* Coefficients



(b) t-test

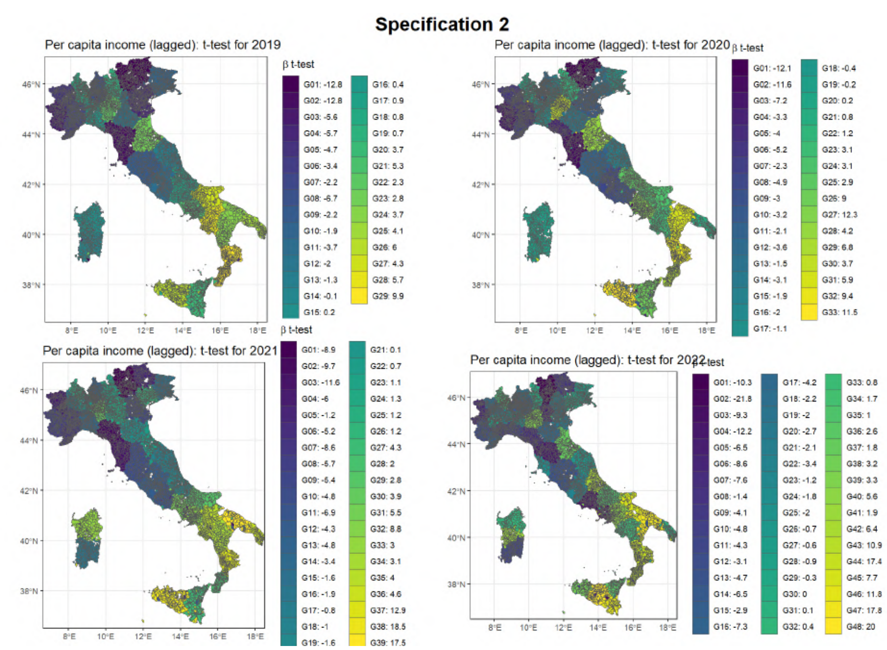
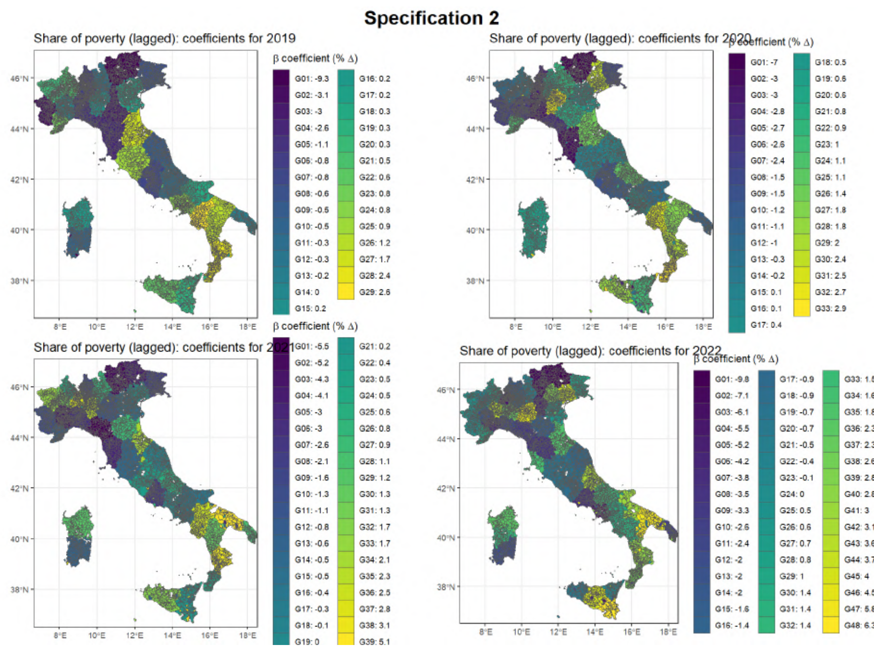


Figure C2: Regression Analysis - *Specification 2*: Results of Share of Poverty by Year of Operation of the GMI

(a) *beta* Coefficients



(b) t-test

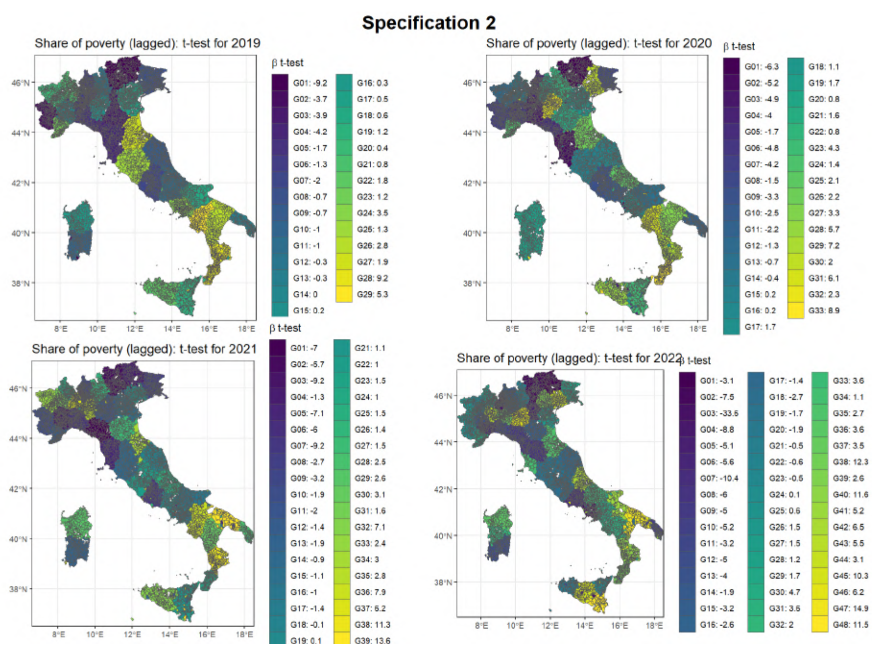
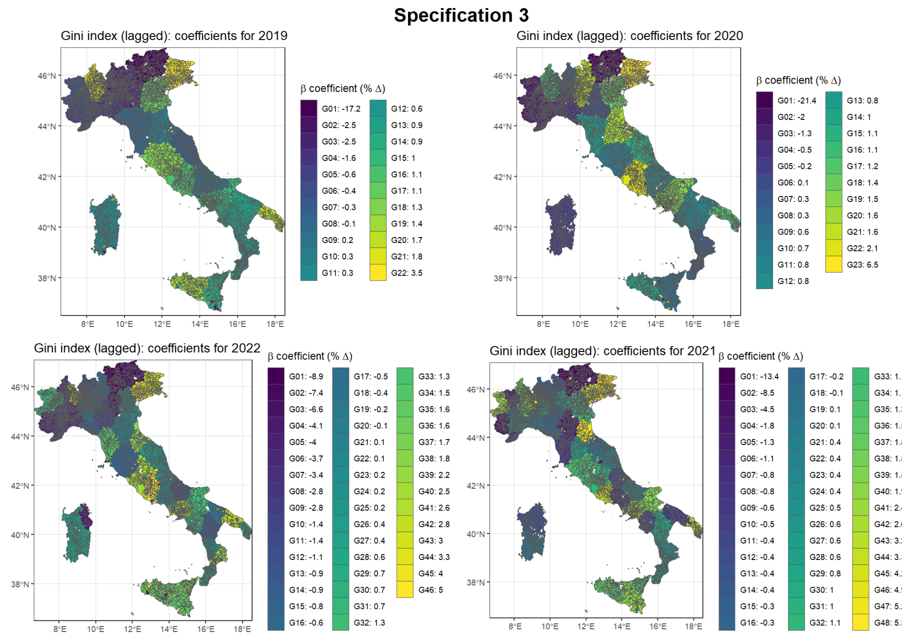


Figure C3: Regression Analysis - *Specification 3*: Results of Gini Index by Year of Operation of the GMI

(a) *beta* Coefficients



(b) t-test

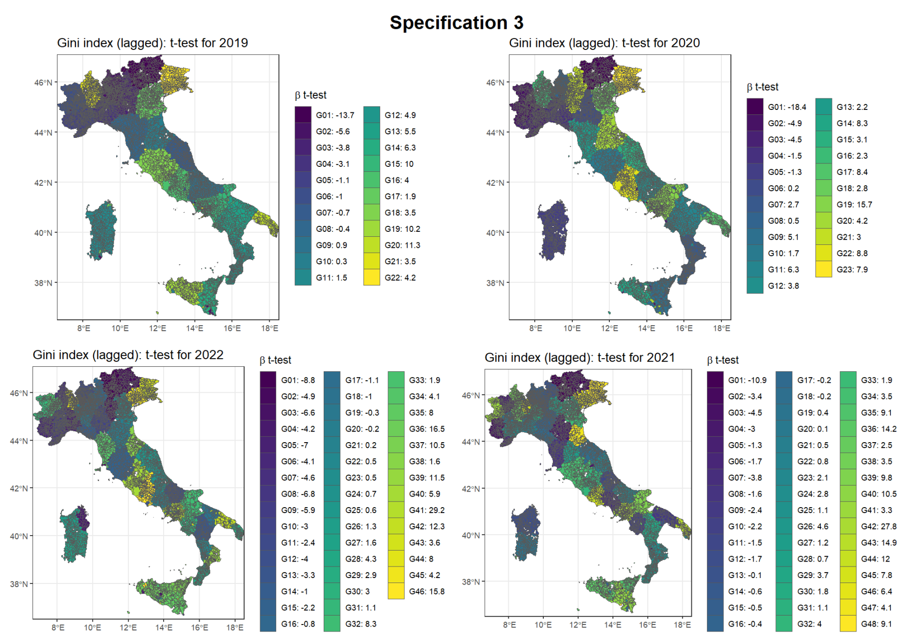
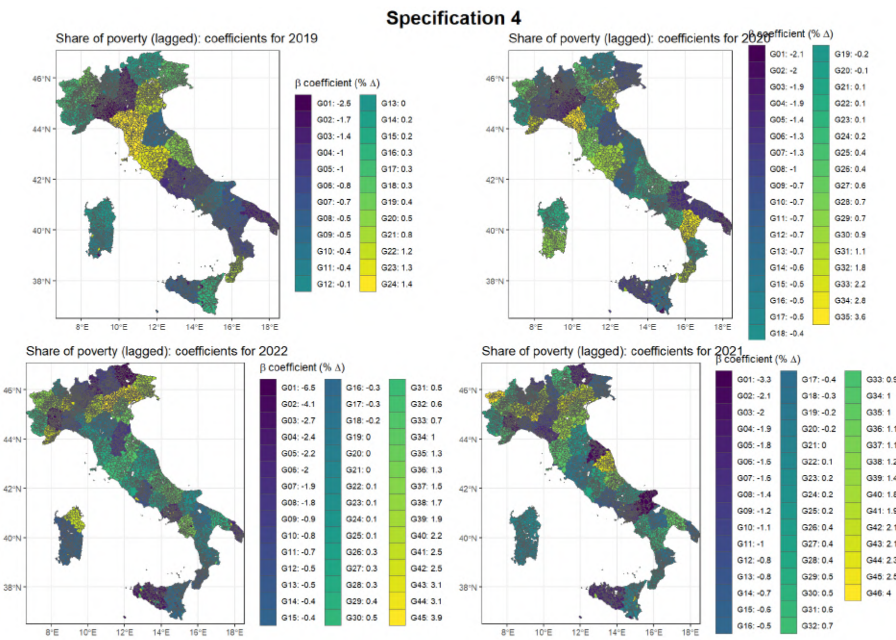


Figure C4: Regression Analysis - *Specification 4*: Results of Share of Poverty by Year of Operation of the GMI

(a) *beta* Coefficients



(b) t-test

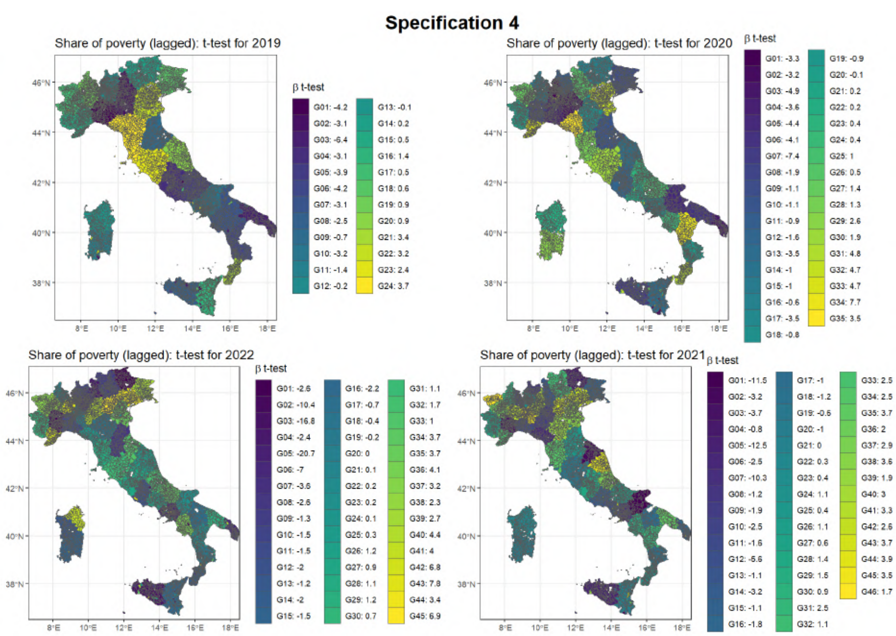
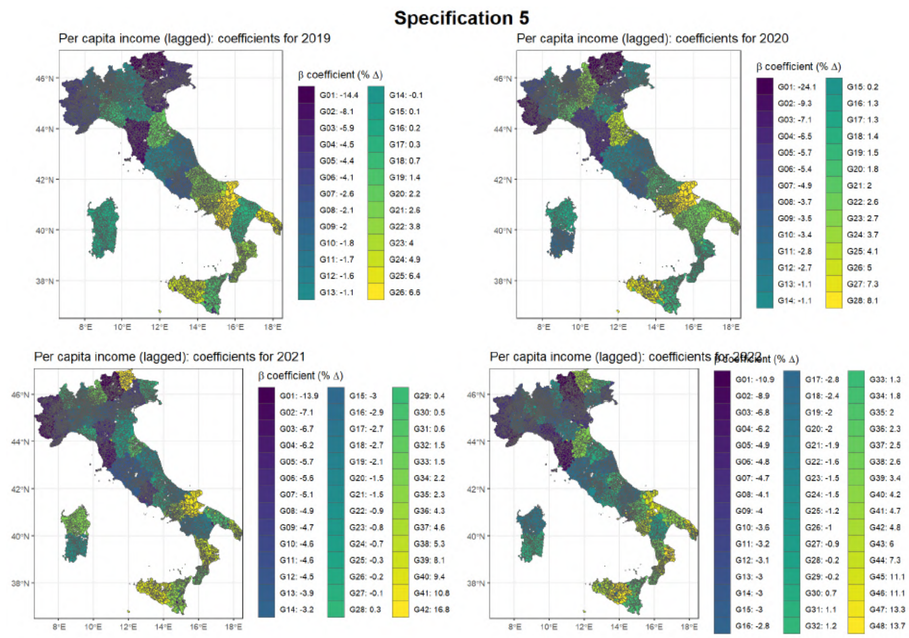
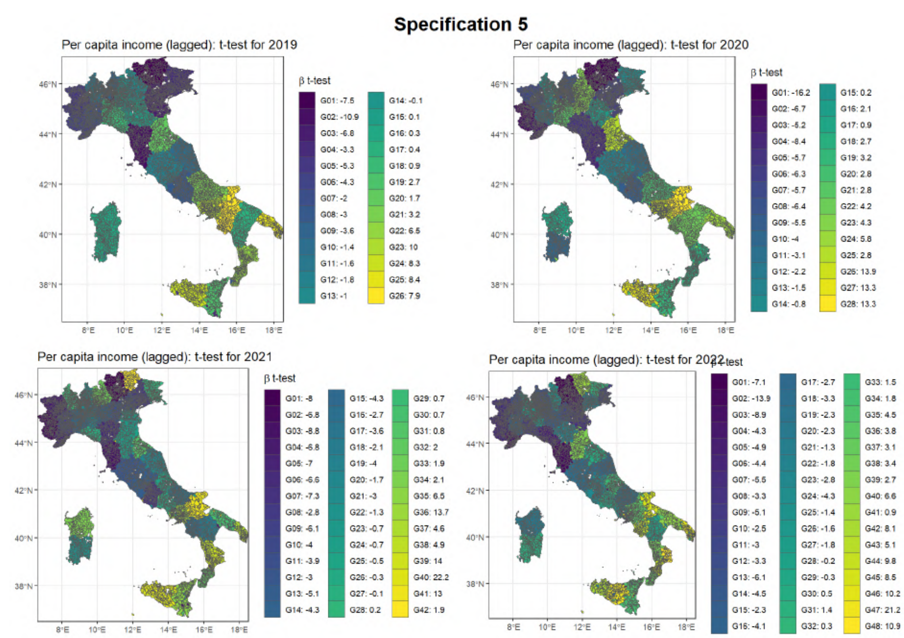


Figure C5: Regression Analysis - *Specification 5*: Results of Per Capita Income by Year of Operation of the GMI

(a) *beta* Coefficients



(b) t-test



Bibliography

Aaberge, R. and Brandolini, A. (2015). Multidimensional poverty and inequality. In *Handbook of income distribution*, volume 2, pages 141–216. Elsevier.

Acemoglu, D. and Restrepo, P. (2018). The race between man and machine: Implications of technology for growth, factor shares, and employment. *American economic review*, 108(6):1488–1542.

Acemoglu, D. and Restrepo, P. (2022). Tasks, automation, and the rise in us wage inequality. *Econometrica*, 90(5):1973–2016.

Albanese, G., Barone, G., and De Blasio, G. (2023). The impact of place-based policies on interpersonal income inequality. *Economica*, 90(358):508–530.

Alesina, A., Di Tella, R., and MacCulloch, R. (2004). Inequality and happiness: are europeans and americans different? *Journal of public economics*, 88(9-10):2009–2042.

Almeida, V., De Poli, S., Hernández, A., et al. (2022). *The effectiveness of Minimum Income schemes in the EU*. European Commission, European Commission.

Ansaloni, V., Aprea, M., Raitano, M., De Smedt, L., Gallo, G., Nicaise, I., et al. (2024). Estimating undercoverage and non-take-up of minimum incomes schemes: Methodological issues and two national case studies.

Antulov-Fantulin, N., Lagravinese, R., and Resce, G. (2021). Predicting bankruptcy of local government: A machine learning approach. *Journal of Economic Behavior & Organization*, 183:681–699.

Aprea, M., Gallo, G., and Raitano, M. (2022). Monetary poverty and adequacy of minimum income schemes in seven major eu countries: What linkages? *Stato e mercato*, 42(3):409–443.

Aprea, M., Gallo, G., and Raitano, M. (2024). The large family penalty in italy: Poverty and eligibility to minimum incomes. *International Journal of Social Welfare*.

Åslund, O. and Fredriksson, P. (2009). Peer effects in welfare dependence: quasi-experimental evidence. *Journal of human resources*, 44(3):798–825.

Atkinson, A. B. (2015). *Inequality: What can be done?* Harvard University Press.

Atkinson, A. B. and Bourguignon, F. (2014). *Handbook of income distribution*. Elsevier.

Atkinson, A. B., Guio, A.-C., and Marlier, E. (2015). Monitoring the evolution of income poverty and real incomes over time.

Baldini, M., Busilacchi, G., and Gallo, G. (2018a). Da politiche di reddito minimo a sistemi integrati nel contrasto alla povertà? un'analisi di dieci paesi europei. *La rivista delle politiche sociali*, 2:189–211.

Baldini, M., Casabianca, E. J., Giarda, E., and Lusignoli, L. (2018b). The impact of rei on italian households' income: A micro and macro evaluation. *Politica economica*, 34(2):103–134.

Banerjee, A., Faye, M., Krueger, A., Niehaus, P., and Suri, T. (2023). Universal basic income: Short-term results from a long-term experiment in kenya.

Bargain, O., Immervoll, H., and Viitamäki, H. (2012). No claim, no pain. measuring the non-take-up of social assistance using register data. *The Journal of Economic Inequality*, 10:375–395.

Bauer, T. K. and Dang, R. (2016). Do welfare dependent neighbors matter for individual welfare dependency? *Available at SSRN 2758653*.

Bavaro, M. and Patriarca, F. (2022). Referrals, intergenerational mobility and human capital accumulation. *Economic Modelling*, 110:105811.

Bhargava, S. and Manoli, D. (2015). Psychological frictions and the incomplete take-up of social benefits: Evidence from an irs field experiment. *American Economic Review*, 105(11):3489–3529.

Blanchard, O. and Rodrik, D. (2023). *Combating Inequality: Rethinking Government's Role*. MIT press.

Bonacini, L., Gallo, G., and Scicchitano, S. (2021). Working from home and income inequality: risks of a 'new normal'with covid-19. *Journal of population economics*, 34(1):303–360.

Boscolo, S. and Gallo, G. (2024). The struggle of being poor and claimant: Evidence on the non-take-up of social policies in italy. *Italian Economic Journal*, 10(2):625–648.

Bourguignon, F. and Morrisson, C. (2002). Inequality among world citizens: 1820–1992. *American economic review*, 92(4):727–744.

Bramley, G., Lancaster, S., and Gordon, D. (2000). Benefit take-up and the geography of poverty in scotland. *Regional studies*, 34(6):507–519.

Brunori, P., Maitino, M. L., Ravagli, L., and Sciclone, N. (2021). Distant and different? lockdown and inequalities in italy. *Economia Pubblica*, (2021/2).

Brunsdon, C., Fotheringham, S., and Charlton, M. (1998). Geographically weighted regression. *Journal of the Royal Statistical Society: Series D (The Statistician)*, 47(3):431–443.

Busilacchi, G. and Fabbri, A. (2023). Does the unemployment trap still exist? the case of the italian minimum income scheme. *Social Policy & Administration*.

Calnitsky, D. (2020). The employer response to the guaranteed annual income. *Socio-Economic Review*, 18(2):493–517.

Cantillon, B., Parolin, Z., and Collado, D. (2020). A glass ceiling on poverty reduction? an empirical investigation into the structural constraints on minimum income protections. *Journal of European Social Policy*, 30(2):129–143.

Card, D., Johnston, A., Leung, P., Mas, A., and Pei, Z. (2015). The effect of unemployment benefits on the duration of unemployment insurance receipt: New evidence from a regression kink design in missouri, 2003–2013. *American Economic Review*, 105(5):126–130.

Card, D., Kluve, J., and Weber, A. (2018). What works? a meta analysis of recent active labor market program evaluations. *Journal of the European Economic Association*, 16(3):894–931.

Celi, G., Ginzburg, A., Guarascio, D., and Simonazzi, A. (2017). *Crisis in the European Monetary Union: A core-periphery perspective*. Routledge.

Cerqua, A. and Letta, M. (2022). Local inequalities of the covid-19 crisis. *Regional science and urban economics*, 92:103752.

Chancel, L. and Piketty, T. (2021). Global income inequality, 1820–2020: the persistence and mutation of extreme inequality. *Journal of the European Economic Association*, 19(6):3025–3062.

Checchi, D. and Peragine, V. (2010). Inequality of opportunity in italy. *The Journal of Economic Inequality*, 8:429–450.

Chetty, R., Friedman, J. N., and Saez, E. (2013). Using differences in knowledge across neighborhoods to uncover the impacts of the eitc on earnings. *American Economic Review*, 103(7):2683–2721.

Conesa, J. C., Li, B., and Li, Q. (2023). A quantitative evaluation of universal basic income. *Journal of Public Economics*, 223:104881.

Connolly, K., Eiser, D., Kumar, A., McGregor, P. G., and Roy, G. (2024). A micro-macro-economic modelling approach to major welfare system reforms: The case of a universal basic income for scotland. *Structural Change and Economic Dynamics*, 68:259–268.

Curci, N., Grasso, G., Savegnago, M., and Recchia, P. (2020). Anti-poverty measures in italy: a microsimulation analysis. *Bank of Italy Temi di Discussione (Working Paper) No*, 1298.

Daruich, D. and Fernández, R. (2024). Universal basic income: A dynamic assessment. *American Economic Review*, 114(1):38–88.

Di Mari, R., Rocci, R., and Gattone, S. A. (2023). Lasso–penalized clusterwise linear regression modelling: a two–step approach. *Journal of Statistical Computation and Simulation*, 93(18):3235–3258.

Eurostat, . (2023). Nuts - nomenclature of territorial units for statistics.

Fadda, S. and Tridico, P. (2017). *Inequality and uneven development in the post-crisis World*. Taylor & Francis.

Feinberg, R. M. and Kuehn, D. (2020). Does a guaranteed basic income encourage entrepreneurship? evidence from alaska. *Review of Industrial Organization*, 57(3):607–626.

Fotheringham, A. S., Yu, H., Wolf, L. J., Oshan, T. M., and Li, Z. (2022). On the notion of 'bandwidth' in geographically weighted regression models of spatially varying processes. *International Journal of Geographical Information Science*, 36(8):1485–1502.

Fox, J. (2015). *Applied regression analysis and generalized linear models*. Sage Publications.

Franzini, M., Patriarca, F., and Raitano, M. (2020). Market competition and parental background wage premium: the role of human and relational capital. *The Journal of Economic Inequality*, 18:291–317.

Frazer, H. and Marlier, E. (2016). Minimum income schemes in europe. a study of national policies 2015. *European Commission*.

Gallo, G. (2021). Regional support for the national government: Joint effects of minimum income schemes in italy. *Italian Economic Journal*, 7(1):149–185.

Gallo, G. and Raitano, M. (2023). Sos incomes: Simulated effects of covid-19 and emergency benefits on individual and household income distribution in italy. *Journal of European Social Policy*, 33(1):101–116.

Gregory, T., Salomons, A., and Zierahn, U. (2022). Racing with or against the machine? evidence on the role of trade in europe. *Journal of the European Economic Association*, 20(2):869–906.

Groot, L., Muffels, R., and Verlaat, T. (2019). Welfare states' social investment strategies and the emergence of dutch experiments on a minimum income guarantee. *Social Policy and Society*, 18(2):277–287.

Grossman, D. and Khalil, U. (2020). Neighborhood networks and program participation. *Journal of Health Economics*, 70:102257.

Gubits, D., Shinn, M., Wood, M., Brown, S. R., Dastrup, S. R., and Bell, S. H. (2018). What interventions work best for families who experience homelessness? impact estimates from the family options study. *Journal of Policy Analysis and Management*, 37(4):835–866.

Guimarães, L. and Lourenço, D. (2024). The imperfections of conditional programs and the case for universal basic income. Technical report, University Library of Munich, Germany.

Hemerijck, A. (2017). *The uses of social investment*. Oxford University Press, Oxford.

Hernandez, M., Pudney, S., and Hancock, R. (2007). The welfare cost of means-testing: pensioner participation in income support. *Journal of Applied Econometrics*, 22(3):581–598.

Iammarino, S., Rodriguez-Pose, A., and Storper, M. (2019). Regional inequality in europe: evidence, theory and policy implications. *Journal of economic geography*, 19(2):273–298.

Immervoll, H., Jenkins, S. P., and Königs, S. (2015). Are recipients of social assistance'benefit dependent'? concepts, measurement and results for selected countries.

Immervoll, H. and Scarpetta, S. (2012). Activation and employment support policies in oecd countries. an overview of current approaches. *IZA Journal of Labor Policy*, 1:1–20.

INPS (2022a). Conoscere il paese per costruire il futuro. XXI Rapporrto Annuale. INPS, Roma.

INPS (2022b). XXI Rapporto Annuale: Relazione del Presidente. INPS, Roma.

INPS (2023). XXII Rapporto annuale. INPS, Roma.

ISTAT, . (2024). Gli indicatori del sistema informativo "a misura di comune". Report.

Jenkins, S. P., Brandolini, A., Micklewright, J., and Nolan, B. (2012). *The Great Recession and the Distribution of Household Income*. OUP Oxford, Oxford, UK.

Kopczewska, K. (2020). *Applied spatial statistics and econometrics: data analysis in R*. Routledge.

Kopczewska, K. (2022). Spatial machine learning: new opportunities for regional science. *The Annals of Regional Science*, 68(3):713–755.

Lelo, K., Monni, S., and Tomassi, F. (2019). Socio-spatial inequalities and urban transformation. the case of rome districts. *Socio-Economic Planning Sciences*, 68:100696.

Liberati, P. (2015). The world distribution of income and its inequality, 1970–2009. *Review of Income and Wealth*, 61(2):248–273.

Liberati, P. and Resce, G. (2022). Regional well-being and its inequality in the oecd member countries. *The Journal of Economic Inequality*, 20(3):671–700.

Locks, G. and Thuilliez, J. (2023). The impact of minimum income on homelessness: Evidence from france. *Journal of Urban Economics*, 135:103547.

Maitino, M. L., Mariani, M., Patacchini, V., Ravagli, L., and Sciclone, N. (2024). The employment effects of the italian minimum guaranteed income scheme reddito di cittadinanza. *Italian Economic Journal*, pages 1–33.

Maket, I. J., Kano, I. S., and Vas, Z. B. (2023). Reverse causal nexus between pro-poor policies and income inequality in kenya. *Regional Science Policy & Practice*, 15(6):1163–1182.

Markussen, S. and Røed, K. (2015). Social insurance networks. *Journal of Human resources*, 50(4):1081–1113.

Marx, I. and Nelson, K. (2013). *Minimum income protection in flux*. Springer.

Milanovic, B. (2016). *Global inequality: A new approach for the age of globalization*. Harvard University Press, Harvard.

Moll, B., Rachel, L., and Restrepo, P. (2022). Uneven growth: automation's impact on income and wealth inequality. *Econometrica*, 90(6):2645–2683.

Monturano, G. (2023). *Comunicare e percepire le politiche economiche e fiscali. L'asimmetria informativa e comunicativa nelle misure di contrasto alla povertà e nelle imposte sui patrimoni*, pages 257–269.

Monturano, G., Resce, G., and Ventura, M. (2023a). The impact of financial support to inner areas on depopulation and industrial growth: A staggered difference-in-difference analysis with spatial spillover effects. *Available at SSRN 4609228*.

Monturano, G., Sonzogno, G. V., and Resce, G. (2023b). La capacità di target territoriale del reddito di cittadinanza. *Menabò. Etica ed Economia*.

Mussida, C. and Parisi, M. L. (2019). Features of personal income inequality before and during the crisis: An analysis of italian regions. *Regional studies*.

Natili, M. (2020). Worlds of last-resort safety nets? a proposed typology of minimum income schemes in europe. *Journal of International and Comparative Social Policy*, 36(1):57–75.

Nifo, A. and Vecchione, G. (2014). Do institutions play a role in skilled migration? the case of italy. *Regional Studies*, 48(10):1628–1649.

Nijman, J. and Wei, Y. D. (2020). Urban inequalities in the 21st century economy. *Applied Geography*, 117:102188.

Nolan, B., Richiardi, M. G., and Valenzuela, L. (2019). The drivers of income inequality in rich countries. *Journal of Economic Surveys*, 33(4):1285–1324.

Palomino, J. C., Rodríguez, J. G., and Sebastian, R. (2020). Wage inequality and poverty effects of lockdown and social distancing in europe. *European economic review*, 129:103564.

Pignataro, G. (2021). Spatial inequality: a multidimensional perspective. *Spatial Economics Volume II: Applications*, pages 123–155.

Piketty, T. and Saez, E. (2014). Inequality in the long run. *Science*, 344(6186):838–843.

Potts, R. B. (1952). Some generalized order-disorder transformations. *Mathematical Proceedings of the Cambridge Philosophical Society*, 48(1):106–109.

Raitano, M. and Vona, F. (2015). Measuring the link between intergenerational occupational mobility and earnings: evidence from eight european countries. *The Journal of Economic Inequality*, 13:83–102.

Rauh, C. and Santos, M. R. (2022). How do transfers and universal basic income impact the labor market and inequality?

Ravallion, M. (2018). Inequality and globalization: A review essay. *Journal of Economic Literature*, 56(2):620–642.

Resce, G. (2022). The impact of political and non-political officials on the financial management of local governments. *Journal of Policy Modeling*, 44(5):943–962.

Rodríguez-Pose, A. (2013). Do institutions matter for regional development? *Regional studies*, 47(7):1034–1047.

Rodríguez-Pose, A. and Hardy, D. (2015). Addressing poverty and inequality in the rural economy from a global perspective. *Applied Geography*, 61:11–23.

Salvati, L. and Carlucci, M. (2014). A composite index of sustainable development at the local scale: Italy as a case study. *Ecological Indicators*, 43:162–171.

Simonazzi, A., Ginzburg, A., and Nocella, G. (2013). Economic relations between germany and southern europe. *Cambridge Journal of Economics*, 37(3):653–675.

Spies-Butcher, B. (2020). Advancing universalism in neoliberal times? basic income, workfare and the politics of conditionality. *Critical Sociology*, 46(4-5):589–603.

Stiglitz, J. E. (2012). *The price of inequality: How today's divided society endangers our future*. WW Norton & Company.

Sugasawa, S. and Murakami, D. (2021). Spatially clustered regression. *Spatial Statistics*, 44:100525.

Terracol, A. (2009). Guaranteed minimum income and unemployment duration in france. *Labour economics*, 16(2):171–182.

Tonutti, G., Bertarelli, G., Giusti, C., and Pratesi, M. (2022). Disaggregation of poverty indicators by small area methods for assessing the targeting of the "reddito di cittadinanza" national policy in italy. *Socio-Economic Planning Sciences*, 82:101327.

Tridico, P. (2018). The determinants of income inequality in oecd countries. *Cambridge Journal of Economics*, 42(4):1009–1042.

Van Parijs, P. and Vanderborght, Y. (2017). *Basic income: A radical proposal for a free society and a sane economy*. Harvard University Press, Harvard.

Wagle, U. R. (2017). How much do social protections matter to poverty and inequality? an insight from asian experiences. *Global social policy*, 17(2):137–167.

Wang, Y., Finazzi, F., and Fassò, A. (2021). D-stem v2: A software for modeling functional spatio-temporal data. *Journal of Statistical Software*, 99(10):1 – 29.

Wheeler, D. C. (2009). Simultaneous coefficient penalization and model selection in geographically weighted regression: The geographically weighted lasso. *Environment and Planning A: Economy and Space*, 41(3):722–742.

Wood, S. N. (2020). Inference and computation with generalized additive models and their extensions. *TEST*, 29(2):307–339.

Wu, B., Li, R., and Huang, B. (2014). A geographically and temporally weighted autoregressive model with application to housing prices. *International Journal of Geographical Information Science*, 28(5):1186–1204.

Xu, Y. and Carraro, L. (2017). Minimum income programme and welfare dependency in china. *International Journal of Social Welfare*, 26(2):141–150.

Yu, L.-r. and Li, X.-y. (2021). The effects of social security expenditure on reducing income inequality and rural poverty in china. *Journal of Integrative Agriculture*, 20(4):1060–1067.

Zhong, S., Wang, M., Zhu, Y., Chen, Z., and Huang, X. (2022). Urban expansion and the urban–rural income gap: Empirical evidence from china. *Cities*, 129:103831.

Zhu, A. X. and Turner, M. (2022). How is the third law of geography different? *Annals of GIS*, 28(1):57–67.

Chapter 4

Exploiting Territorial Disparities to Evaluate the Impact of Income Support on Electoral Behavior in Italy*

Abstract

In this study, we explore the relationship between government income support programmes and voters' behaviour, leveraged by the Italian Citizenship Income measure. We create a unique dataset by merging administrative data on programme beneficiaries with electoral data at the municipal level. Through a Difference-in-Difference approach with continuous treatment that exploits the variation in the share of beneficiary families, the research reveals that the income support programme affects citizens' voting behaviour in favour of the ruling party, but only in disadvantaged regions. In contrasting circumstances, the income support policy has a contrary impact on voting behaviour. The results emphasise the importance of taking into account regional differences and contextual elements when assessing how income support programmes influence political results.

JEL classifications: H53; D72; C21; C33; I38

Keywords: Income Support Policies; Pocketbook Voting; Sociotropic Voting; Italian Citizenship Income; Difference-in-Difference (DiD); Political Economy

---

*This work is a paper co-authored with **Professor Giuliano Resce** (Department of Economics, Via Francesco De Sanctis, 86100 Campobasso CB, and International Inequalities Institute, London School of Economics and Political Science, Houghton Street, London, WC2A 2AE, UK) and **Dr. Giulia Valeria Sonzogno** (Department of Economics, Roma Tre University, Via Silvio D'Amico, 77, 00145 Rome, Italy). We thank the Italian Institute of Social Security (INPS) for providing the data.

4.1 Introduction

One of the most important areas of debate in political economy, social choice and public economics revolve around the link between public policies and voters' behaviour. The theory of democracy is built on the idea that voters react to policy outcomes, forming the basis for government accountability through this mechanism (Lewis-Beck, 1985; Manacorda et al., 2011). However, although the outcomes of implemented policies garnered significant attention over time, their impact on voting behaviour remains largely uncertain (Lewis-Beck and Stegmaier, 2000, 2008, 2018; Elinder et al., 2015).

A prominent area of research has focused on examining whether specific government initiatives have the potential to boost support for the party currently in power, and a consensus has emerged in many studies regarding the role of government welfare programmes. These programmes provide incumbents with valuable opportunities to take credit for their positive outcomes, as noted by Pierson (1996). Well-targeted allocations of public resources can sway the electorate in support of the ruling party, and it seems plausible that government programmes, tailored to address specific population segments, could potentially increase backing for the incumbent party. However, the empirical evidence on this issue remains rather scant. Prior research has predominantly concentrated on Latin American developing nations (Cerda and Vergara, 2008; Manacorda et al., 2011; De La O, 2013; Zucco Jr, 2013; Conover et al., 2020; Araújo, 2021). These investigations consistently report favourable impacts of cash transfers on backing for ruling parties. Similarly, Pop-Eleches et al. (2012) identified parallel effects related to cash transfers in Romania using stated preferences data, and more recently, Gromadzki et al. (2024) demonstrated that cash transfers resulted in an increased vote share for the ruling party in Poland. A notable gap in this literature is the consideration of contextual variables that could potentially moderate the relationship between tailored government initiatives and political outcomes, encompassing factors connected to cultural variations, historical backgrounds and economic circumstances.

This article explores the role of regional disparities in the connection between income support programmes and voting for the ruling party. To this purpose, we leverage the Italian Citizenship Income (Italian: *Reddito di cittadinanza*, hereafter also CI), a substantial anti-poverty initiative introduced in 2019. To test the effect of the policy on the support for the party that proposed the programme, we have built a unique dataset by merging administrative data on the share of households receiving the CI with electoral data at the municipal level. The effect is estimated by exploiting the municipal heterogeneity in the share of beneficiaries' households which allows us to implement a generalised Difference-in-Difference approach with continuous treatment, using the share of votes taken by the party that invented and supported the policy (Five Stars Movement, hereafter also IP - Incumbent Party) as the outcome variable.

Italy is a strong candidate for investigating the role of contextual variables because, among the European countries, it is the one with the highest level of economic inequality among its regions (Eurostat, 2019). The Northern regions compete with the Centre of Europe, notably Germany, while the Southern regions align more closely with Greece (Chelli et al., 2023). Results show that the income support programme affects citizens' voting behaviour in favour of the ruling party only in less developed regions, where the combination of lower institutional quality and high unemployment undermines prospects for development. In such contexts, income support emerges as a viable mechanism for bolstering support for the incumbent party, whereas in contrasting circumstances, the income support policy has a contrary impact on voting behaviour.

The remainder of the document is structured as follows: Section 4.2 explains the institutional context and the potential mechanisms; Section 4.3 presents the empirical strategy and the data; Section 4.4 shows the results of the econometric estimates; Section 4.5 tests the heterogeneity of the results and Section 4.7 concludes.

4.2 Institutional and Theoretical Framework

4.2.1 The Italian Citizenship Income

During the last decades, the level of income inequality between and within countries has increased, taking central stage in the political and academic debate worldwide (Alesina and Perotti, 1996; Piketty and Saez, 2014; Atkinson, 2015). Consequently, reducing inequalities and promoting upward convergence in living conditions have become pressing needs, on top of political agendas. To face these challenges, variegated tools and solutions have been studied and implemented around the world (Atkinson and Piketty, 2007; Arbia, 2023; Blanchard and Rodrik, 2023). Introduced in 2019 the CI (Italian Citizenship Income)[1] represents the first large-scale citizens' income program introduced in the country after experiencing similar partial measures (Baldini and Gori, 2019; Gallo, 2021; Busilacchi and Fabbri, 2023; Monturano, 2023; Monturano et al., 2023; Maitino et al., 2024).

The CI, illustrated largely in detail in Chapter 3, provided economic support to the income of families combined with a program of reintegration into work and social inclusion (Baldini and Gori, 2019; Busilacchi and Fabbri, 2023; Monturano, 2023; Maitino et al., 2024). Income support has been available to Italian families (citizens resident in Italy for at least 10 years) since March 2019. The measure is based on income and residence requirements: the amount of the benefit is determined by the age of family members and the ownership of the house. Families can apply if their income and assets are below a certain threshold and, after verification, the benefit is paid for up to 18 months and can be renewed.

[1]Decree Law of 28 January 2019, no. 4, converted with Law of 28 March 2019, no. 26.

During the entire period of application of the measure, April 2019 - December 2023[2], the Italian Citizenship Income has reached a large number of households (INPS, 2023). In the first months of application (between April and December 2019), the CI had already reached almost 1 million families and more than 2 million individuals[3].

Since 2020, there has been an increase in the number of beneficiary households (1.4 million families and over 3.5 million individuals), reaching a peak of 1.6 million families and over 3.7 million citizens in 2021. The growth in the number of families treated is largely attributable to the induced effect of COVID-19 which has worsened the socio-economic condition of Italian citizens (Palomino et al., 2020; Carta and De Philippis, 2021; Raitano et al., 2021). In this regard, recent studies have shed light on its role during the pandemic crisis (INPS, 2022b,a; Gallo and Raitano, 2023).

With the reduction of restrictions and the reopening of markets in 2022, a significant decrease in the number of beneficiaries of the measure has been observed. The reduction in the number of households treated was then further accentuated in 2023 (with just over 1.2 million and around 2.7 million individuals), influenced by the legislative changes made to the Law, which restricted the eligibility criteria for obtaining the subsidy[4].

Table 4.1 presents the descriptive statistics of our CI dataset, focusing on beneficiary families at regional and macro-regional levels (i.e. at NUTS 1 and 2 levels). The data covers a three-year period, from 2019 to 2021, which corresponds to the interval between two Italian national political elections (March 2018 and September 2022). The intention to approve a national income support measure was first announced by the party during the electoral campaign of February 2013, when it presented itself for the first time in the general political elections, and subsequently through the launch of actions and legislative initiatives in Parliament[5]. However, it is only during the 2018 electoral campaign that the

[2]The Italian Citizenship Income, introduced in 2019, was subjected to amendments. The 2023 Budget Law, Law No. 197 of 29 December 2022, restricted, starting from September 2023, the provision of the subsidy only to families with disabled people or people over sixty and at the same time established the abolition starting from January 2024. Subsequently, with the Decree of 4 May 2023, No. 48, converted into law on 3 July 2023, No. 85, the Italian Government introduced two new subsidies to replace the CI: 'the Inclusion Allowance' and the 'Support for training and work'.

[3]The reported values are present in the Quarterly Income Report of Citizenship managed by the Statistical Observatory of the National Institute of Social Security (INPS) and refer to all families and individuals who have received at least one month's CI. The Report is available at this address: https://www.inps.it/it/it/dati-e-bilanci/osservatori-statistici-e-altre-statistiche/dati-cartacei---rdc.html.

[4]The legislative changes to the Italian Citizenship Income introduced with Law No. 197 of 29 December 2022 include the limitation of the provision of the subsidy only to specific categories, as well as the introduction of new replacement measures, to determine a reorientation of policies of welfare in the Italian context.

[5]Draft law presented to the Senate on 29 October 2013, concerning "Establishment of citizen's income as well as a delegation to the Government for the introduction of the minimum hourly wage". Available at https://www.senato.it/leg/17/BGT/Schede/Ddliter/42593.htm.

IP made explicit the possibility of implementing a universalistic-selective support measure. The measure was implemented only in 2019, when the party came into government, and it remained unchanged until the first regulatory change in 2022. Consequently, based on this, the period of the CI analyzed in our work offers a complete perspective of the potential impacts of the measure on electoral outcomes.

Table 4.1 highlights an increasing trend in households receiving the measure along the three years of observation. The geographical distribution of beneficiaries reproduces the historical Italian socioeconomic divide (Acciari and Mocetti, 2013; Felice, 2018; Coco and De Vincenti, 2020; Accetturo et al., 2022). Southern regions account for the highest incidence of families receiving CI compared to the rest of Italy. In 2021, across South and Islands regions, the number of beneficiary households out of total households exceeded, on average, 15 percent, reaching the value of 10 percent on a three-year basis. In the same year, the peak was reached in Campania (20 percent), Sicily (19 percent) and Calabria (17 percent).

In contrast, Central regions had lower values of beneficiary households (4 percent on a three-year basis). In Northern regions, the share of beneficiary households was significantly smaller than in South and Central Italy. In the three years of implementation of the policy in the North-East, less than 2 percent of families benefited from the measure. Similar values are recorded in the North-West with just 3 percent of families receiving CI. In 2021, Trentino-Alto Adige was the region with the lowest percentage of beneficiary households (1.61 percent).

Table 4.1: Households Receiving Italian Citizenship Income (CI)

Territorial	**Number of households receiving CI**			**Regional percentage of CI households on total households by region**		
	2019	**2020**	**2021**	**2019**	**2020**	**2021**
Abruzzo	18197	23522	40712	3.3	4.22	7.29
Basilicata	9176	10657	17083	3.9	4.49	7.2
Calabria	58736	80038	136188	7.37	9.96	16.85
Campania	160850	252257	455676	7.44	11.46	20.59
Emilia-Romagna	27344	36772	68471	1.36	1.81	3.37
Friuli-Venezia Giulia	8979	11236	18432	1.6	1.98	3.26
Lazio	73481	110539	221167	2.84	4.2	8.41
Liguria	17130	24286	43756	2.26	3.17	5.74
Lombardy	65070	95093	184521	1.46	2.11	4.11
Marche	11999	15682	26923	1.88	2.42	4.16
Molise	5076	6809	11238	3.9	5.21	8.59
Piemonte	46125	64059	117005	2.31	3.19	5.84
Puglia	78324	111801	194572	4.91	6.88	11.89
Sardinia	38044	47711	79643	5.24	6.49	10.78
Sicily	148161	222003	388026	7.39	10.83	18.78
Tuscany	30946	39533	70235	1.89	2.36	4.22
Trentino-Alto Adige	2414	3504	7584	0.53	0.74	1.61
Umbria	8636	11748	20486	2.27	3.05	5.34
Aosta Valley	908	1068	1812	1.51	1.76	3
Veneto	23918	31011	56012	1.15	1.47	2.66
Northeast	62655	82523	150499	1.22	1.59	2.91
Northwest	129233	184506	347094	1.78	2.51	4.74
Central	125062	177502	338811	2.38	3.33	6.36
South	330359	485084	855469	6.04	8.73	15.32
Islands	186205	269714	467669	6.82	9.68	16.67
Italy	833514	1199329	2159542	3.23	4.58	8.24

Source: Author's processing of INPS data.

4.2.2 Citizens' Income Support Measure and Electoral Results

Long-term effects of the Italian CI on the labour market, productivity, and other economic and demographic variables have not been fully explored to date. Recent studies have explored the effect of CI in addressing the greatest poverty, vulnerability and social exclusion, especially during the Covid-19 pandemic, with mixed results (Checchi et al., 2021; Gallo, 2021; Tonutti et al., 2022; Busilacchi and Fabbri, 2023; Gallo and Raitano, 2023; Maitino et al., 2024). The link between CI and electoral results is still unexplored. In this sense, votes obtained by the IP in the political elections for the renewal of Parliament in the elections in which the approval of the CI was discussed (i.e., 2013, 2018, and 2022), represent a unique case study to investigate whether the introduction of welfare measures generates electoral returns.

It has been shown that electoral outcomes in Italy and Western nations are influenced by factors like socioeconomic disparities, globalisation, multidimensional poverty, labour market instability, institutional deficiencies, political volatility and economic policies. For example, Alesina and Fuchs-Schündeln (2007) focusing on individual preferences for redistribution and the generosity of welfare systems, highlight how these differ significantly between countries due to government systems and other political-institutional dynamics. Rodríguez-Pose (2018) demonstrates that persistent poverty, economic decline and the lack of job opportunities are at the roots of electoral heterogeneity between different territorial areas, shedding light on the increasing anti-establishment vote in stagnant regions and in the so-called 'left-behind places'. In another study, Rodríguez-Pose et al. (2021) explore these issues further, in the context of the United States. They highlight how areas with strong social capital, but experiencing economic and demographic decline, were more likely to support anti-establishment candidates, suggesting that the discontent behind the populist wave is deeply rooted in long-term economic distress. Similar results emerge from the work of Dijkstra et al. (2020), a study of over 63,000 electoral districts in European countries, which demonstrates that long-term economic problems, lack of industrial policies, scarcity of job opportunities, and low levels of human capital in the labour market contribute to increasing support for Eurosceptic parties. Even in the case of Brexit, economic decline and socioeconomic inequalities were important factors in shaping electoral consensus (Alabrese et al., 2019; Carreras, 2019).

Other works, study the electoral behaviour of precarious workers. Marx (2016) and Rovny and Rovny (2017) show that workers in conditions of greater insecurity and vulnerability are more likely to blame the government in power for their economic situation, to support redistributive measures, and to vote for radical and extremist parties or to abstain. Piketty (2018) and Gethin et al. (2022) study long-term electoral dynamics in various countries focusing on the role of education and income, as socioeconomic factors, capable of influencing election behaviour among different classes of voters.

Imami et al. (2023) analyse the phenomenon of Electoral Politics of Disaster (EPD), i.e., the electoral return deriving from adopting specific government spending measures following natural and socioeconomic disasters. The authors demonstrate, in the context of Albania, that resources for disasters such as COVID-19 and the 2019 earthquake were selectively allocated to specific geographic areas that were more likely to support the ruling party, and that this positively influenced the outcome of the elections, increasing consensus towards the ruling party. More recently, Cerqua et al. (2023), by focusing on two earthquakes in Italy, show that the efficiency of institutions is a crucial element in changes in political consensus following natural disasters. The electoral heterogeneity in Italy also seems to be influenced by the level of EU funding: municipalities with higher levels of funding, experience negative and significant impacts on voting for populist and anti-system parties Albanese et al. (2022).

Place of residence, level of insecurity, unemployment, employment profile of workers, and level of education represent additional elements that drive the behaviour of voters at a territorial level (Levi and Patriarca, 2020). In fact, Pignataro and Prarolo (2020) highlight how spatial proximity and learning influence electoral behaviour. Studying Italian referenda on nuclear energy, the authors observe that proximity to an existing or proposed nuclear site intensifies opposition, especially in the vicinity of proposed sites. In contrast, prolonged exposure to existing sites, combined with a learning process based on increased local information, reduces perceived risk and, consequently, electoral opposition to such plants. However, Caselli et al. (2020) assessed the impact of globalisation on electoral dynamics in Italy between 1994 and 2008 and showed that globalisation had played a significant role in shaping electoral outcomes. Levi et al. (2020), by exploring the effect of immigration on voting for parties with a programme hostile to the reception of migrants, conclude that the electoral expansion of these movements could be a short-term temporary effect.

From these works, it is clear that inequalities between regions in Italy are key factors in describing electoral heterogeneity. The works of Bloise et al. (2021) and Bloise et al. (2023) analysed the influence of the economic and social inequalities in Italy on votes for the national party, exploring the electoral shocks that occurred in Italy between 1994 and 2018. Their results indicate that economic conditions and regional disparities have a significant influence on votes. In particular, Basile (2023) shows how regional, socioeconomic and institutional inequalities negatively affected electoral participation in less developed regions. Institutional factors do not appear to have a significant impact, especially on participation. In this regard, more than the quality of institutions, institutional efficiency seems to influence electoral outcomes (Ippolito et al., 2023). With reference to regional disparities, Albanese et al. (2024) examined the impact on political preferences of an important large state measure aimed at encouraging the development of the Southern regions i.e.'*Cassa per il Mezzogiorno*', implemented in the second half of the twentieth century. The results highlight that, in territories that have benefited from greater state aid, citizens tend to have greater preferences for similar future policies. Consequently, these results suggest that the measure has permanently changed the preferences and perceptions of citizens of the Italian

South on redistribution. They also find that these subsidies resulted in deferred electoral rewards for the IP in our study.

4.2.3 The Evolution of Voting Behaviour Among Three National Elections in Italy

In the 2022 national elections, the party that proposed the CI obtained much lower support than in 2013 and 2018, losing its government position. This decline in support was consistent in all Italian regions and in both elective assemblies, with the national percentage of votes falling to less than 16 percent, compared to 25 percent in 2013 and 33 percent in 2018 (Chiaramonte et al., 2022).

Table 4.2 reports the variation of votes at regional and macro-regional levels for the IP in the general elections by comparing the election years in which the CI was proposed by the IP with the year of voting in which the measure was in force (i.e. 2013-2022 and 2018-2022). Between 2013 and 2022 there was a marked decline in consensus for the IP, spatially concentrated in the Central-Northern regions. In the Southern regions and Islands, the consensus for the IP remained almost stable between the two elections considered. However, in the 2018 and 2022 elections, higher losses of support are recorded in all regions. In these two electoral rounds, the greatest losses are concentrated in the territories where the IP has obtained the highest percentages of votes in the 2018 national elections, i.e. in the Southern and Island regions and along the coastal strip of the central Adriatic (i.e. Marche), with a reduction in consensus of around 14 percent (see Table 4.2 and Figures 4.1, D1). However, in the Northern regions, the loss of consensus was more limited. From this perspective, Trentino-Alto Adige records the smallest variation in votes compared to the total number of voters (-10 percent).

Despite the general decline in votes, as can be seen from Figures 4.1, D1, in the 2022 elections the support for the proposing party increased in many municipalities and in several single-member constituencies. In several cases, the IP overcomes other parties and coalitions, as in the case of the single-member constituencies of the metropolitan city of Naples or some Puglia constituencies (Chiaramonte et al., 2022). The crystallisation of the consensus of the IP in many areas of the South between 2018 and 2022, has been considered the result of the approval of specific welfare and social protection measures, which the same political force first announced and then contributed to its approval[6] during the governments in which they were part of the majority[7] (Angelucci et al., 2022; Giuliani,

[6]For example the ‘Italian Citizenship Income’ (CI) and the increase in pensions for disabled people in economic difficulty, i.e. ‘Citizenship Pension’, the subsidy against the economic emergency of Covid-19 ‘Emergency Income’, the reform of the labour market ‘Dignity Decree’, the reform of the pension system ‘Quota 100’, the Super Bonus for the energy and seismic efficiency of buildings ‘Super-Bonus 110 percent’ etc.

[7]Conte I Government, in coalition with the Northern League; Conte II government with Democratic Party and other centre-left parties and movements; Draghi government, grand coalition.

2023). Among these policies, the most important and influential is the Italian Citizenship Income, approved in 2019.

Table 4.2: Regional Variation in Votes for the Incumbent Party (IP) in General Elections Between 2013-2022 and 2018-2022

Territorial	**2013-2022**				**2018-2022**			
	Absolute regional change in votes of the IP		**Changes regional in voting shares of IP on the number of voters by region**		**Absolute regional change in votes of the IP**		**Changes regional in voting shares of IP on the number of voters by region**	
	Chamber of Deputies	**Senate**	**Chamber of Deputies**	**Senate**	**Chamber of Deputies**	**Senate**	**Chamber of Deputies**	**Senate**
Abruzzo	-121697	-89240	-10.99	-9.6	-177989	-148809	-16.74	-16.5
Basilicata	-16479	-6362	-3.28	-1.55	-73763	-60308	-15.5	-15.18
Calabria	-22624	23610	-1.42	1.85	-178236	-144731	-11.7	-11.51
Campania	118224	235490	2.54	6.5	-654457	-479730	-14.2	-13.08
Emilia-Romagna	-428625	-346836	-13.16	-12.25	-441191	-379402	-13.34	-13.3
Friuli-Venezia Giulia	-152252	-127493	-15.76	-15.13	-115515	-103275	-11.98	-12.35
Lazio	-541581	-402908	-12.01	-10.77	-590971	-507404	-13.33	-13.52
Liguria	-209693	-173634	-15.5	-14.57	-157574	-135299	-12.47	-12.25
Lombardy	-758229	-569284	-10.34	-9.13	-780886	-704671	-10.46	-11.11
Marche	-191842	-152640	-16.1	-14.97	-198325	-172333	-16.63	-16.94
Molise	-21508	-14553	-7.68	-6.23	-43570	-36561	-16.48	-16.36
Piedmont	-495988	-405569	-14.03	-13.24	-399947	-354067	-11.7	-12.02
Puglia	-95823	-7254	-2.84	-0.27	-465917	-362032	-14.07	-13.31
Sardinia	-131608	-100803	-9.16	-8.34	-209481	-182795	-15.03	-15.28
Sicily	-278375	-118236	-6.63	-3.54	-573853	-472366	-14.11	-14.23
Tuscany	-324546	-258732	-11.29	-10.39	-303104	-259306	-10.59	-10.48
Trentino-Alto Adige	-58411	-48748	-8.47	-8.71	-75687	-63334	-9.84	-10.11
Umbria	-89897	-69920	-12.74	-11.51	-80184	-68422	-11.75	-11.63
Aosta Valley	-6791	-7393	-7.29	-9.24	-9757	-8594	-10.07	-10.41
Veneto	-627271	-523923	-17.19	-16.9	-505029	-465313	-13.74	-14.97
Northeast	-1266559	-1047000	-14.79	-14.27	-1137422	-1011324	-13.05	-13.62
Northwest	-1470701	-1155880	-11.94	-10.94	-1348164	-1202631	-11.01	-11.48
Central	-1147866	-884200	-12.37	-11.25	-1172584	-1007465	-12.79	-12.86
South	-159907	141691	-1.39	1.54	-1593932	-1232171	-14.17	-13.44
Islands	-409983	-219039	-7.27	-4.82	-783334	-655161	-14.34	-14.51
Italy	**-4455016**	**-3164428**	**-9.42**	**-8.01**	**-6035436**	**-5108752**	**-12.89**	**-12.96**

Source: Author's processing on Ministry of the Interior data, Eligendo portal.

4.2.4 The CI as a Possible Driver for the Electoral Results of the Incumbent Party

The concept of pocketbook emerges as one of the most significant theories in political economy and the theoretical framework of electoral behaviour. Pocketbook refers to the tendency of voters to evaluate the incumbent government based on how its policies have directly affected their wallets. In this regard, Lewis-Beck (1985) in one of the first and most influential works in this field, showed that voters tended to support the incumbent party when their personal circumstances were financially profitable. In this case, voters' behaviour is positively conditioned by economic promises on transfers, such as resources earmarked for families with children (Elinder et al., 2015) or measures to combat poverty (Manacorda et al., 2011).

In contrast, the concept of sociotropic voting refers to voting behaviour that is influenced by local or global socioeconomic conditions, rather than by individual opportunistic choices. More precisely, voters' perceptions of a country's economic health and/or evaluations of political and economic institutions can significantly influence voting behaviour and thus overall electoral outcomes (Lewis-Beck and Stegmaier, 2000, 2018).

Both of these economic theories fall within a broader concept: retrospective voting. In it, voters decide who to entrust their consent to by analysing and evaluating the economic policies approved by governments in office. In this context, past policies become a weighting factor for making future electoral decisions (Fiorina, 1978).

Some contributions have focused on specific welfare measures, implemented by governments in office, highlighting the growth of electoral support for the ruling party (Pierson, 1996). Other works, such as Levitt and Snyder Jr (1997), Cerda and Vergara (2008), Cox (2009) have instead focused on the study of the allocation of public resources. The analysis of federal government spending in the United States on the electoral outcomes of members of Congress (Levitt and Snyder Jr, 1997) and the study of government subsidies on the presidential election outcomes in Chile (Cerda and Vergara, 2008), highlight the complexity of identifying clear relationships and dependencies between public spending policies and voting behaviour.

In this context, the relationship between income support measures and electoral outcomes represents one of the most debated and understudied topics in the political and academic debate (Lalioti, 2016; Loeffler, 2023). To address this issue, this work uses an evidence-based approach to shed light on how (and if) income support measures influence electoral choices to offer relevant insights on how the adoption of social welfare policies can lead to an increase in votes for the parties that promote them.

According to the theory that considers voters as economic agents maximising their benefits (i.e. the pocketbook vote), the share of families receiving the Italian CI should be a driver for an increase in consensus for the political force that promoted and supported it, i.e. the IP (Lewis-Beck, 1985; Cerda and Vergara, 2008; Manacorda et al., 2011; Elinder et al., 2015). However, the link between local benefits and local support for the IP may be mediated by different levels of social and economic deprivation at the territorial scale, as shown, among others, by Lewis-Beck and Stegmaier (2000, 2008, 2018).

Focusing on this issue, for Girardi (2023), the electoral consensus of the IP derives from problems linked to the labour market and specifically from the precariousness that develops in it. Angelucci et al. (2022), argues that the consensus for IP is instead driven by citizens' distrust of public institutions. Similar results emerge from the work of Faggian et al. (2021), which highlights how the IP has achieved significant results in Southern regions. In these areas, the levels of unemployment, poverty, public sector inefficiency and the corresponding need to address these issues are higher. Giuliani (2023) summarises these different interpretations showing a positive association between the number of beneficiaries of Italian CI and the votes for the incumbent party which turns out to be negative when controlling for confounding factors.

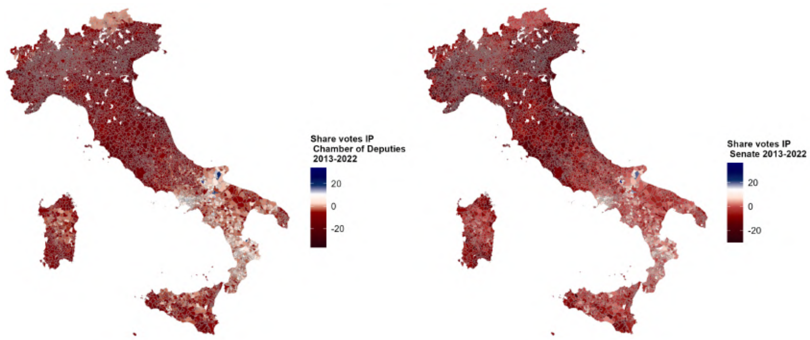
Figures 4.1 and 4.2 show the electoral results for the IP and the share of households receiving CI. Comparing the maps, it seems that the CI has not converted into a direct increase in support for the IP. This evidence contrasts with the widespread idea that direct welfare measures, such as the CI, can act as catalysts for the electoral consensus of the parties that propose them (Lewis-Beck and Stegmaier, 2000, 2008, 2018), indicating, differently, that other factors can prevail more incisively in the decisions and behaviour of voters (Faggian et al., 2021; Angelucci et al., 2022; Girardi, 2023). However, the absence of a direct parallel between welfare benefits and electoral consensus is not uniform throughout Italy. In some areas of the country, and in particular, in specific areas of the South, it is possible to observe a high number of treated households and an increase in support for IP, which suggests a possible favourable link between the CI and electoral consensus due to specific local contexts.

In light of the considerations revealed in this section, the research questions that guide our study and that we intend to answer are the following:

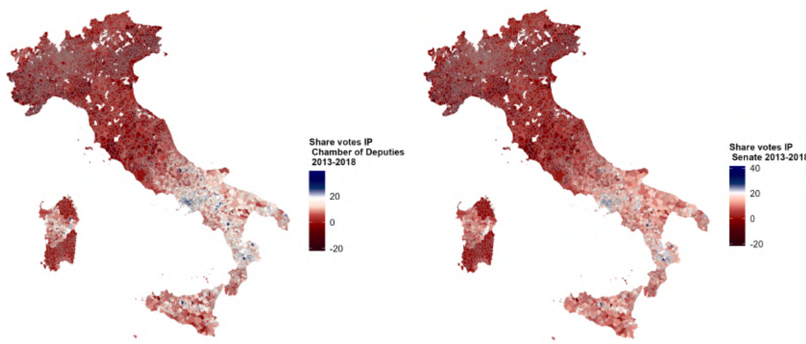
- **Research Question RQ$_1$** - *Has the income support policy an effect on the votes obtained by the political force that promoted it?*
- **Research Question RQ$_2$** - *Is there any mediating role of socioeconomic and institutional heterogeneity?*

Figure 4.1: Change in Votes for the Incumbent Party on the Number of Electors - Chamber of Deputies and Senate (2013-2022)

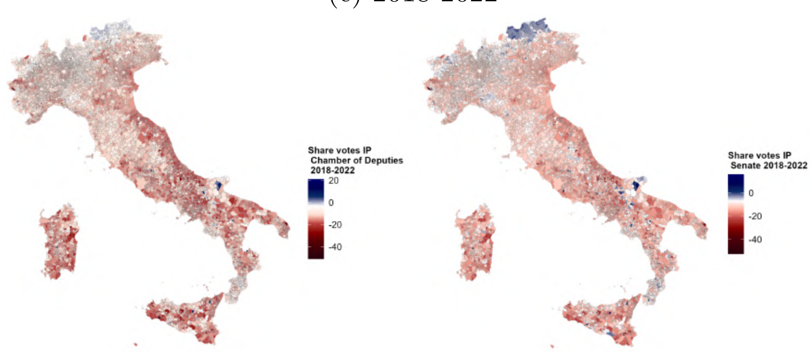
(a) 2013-2018



(b) 2013-2022

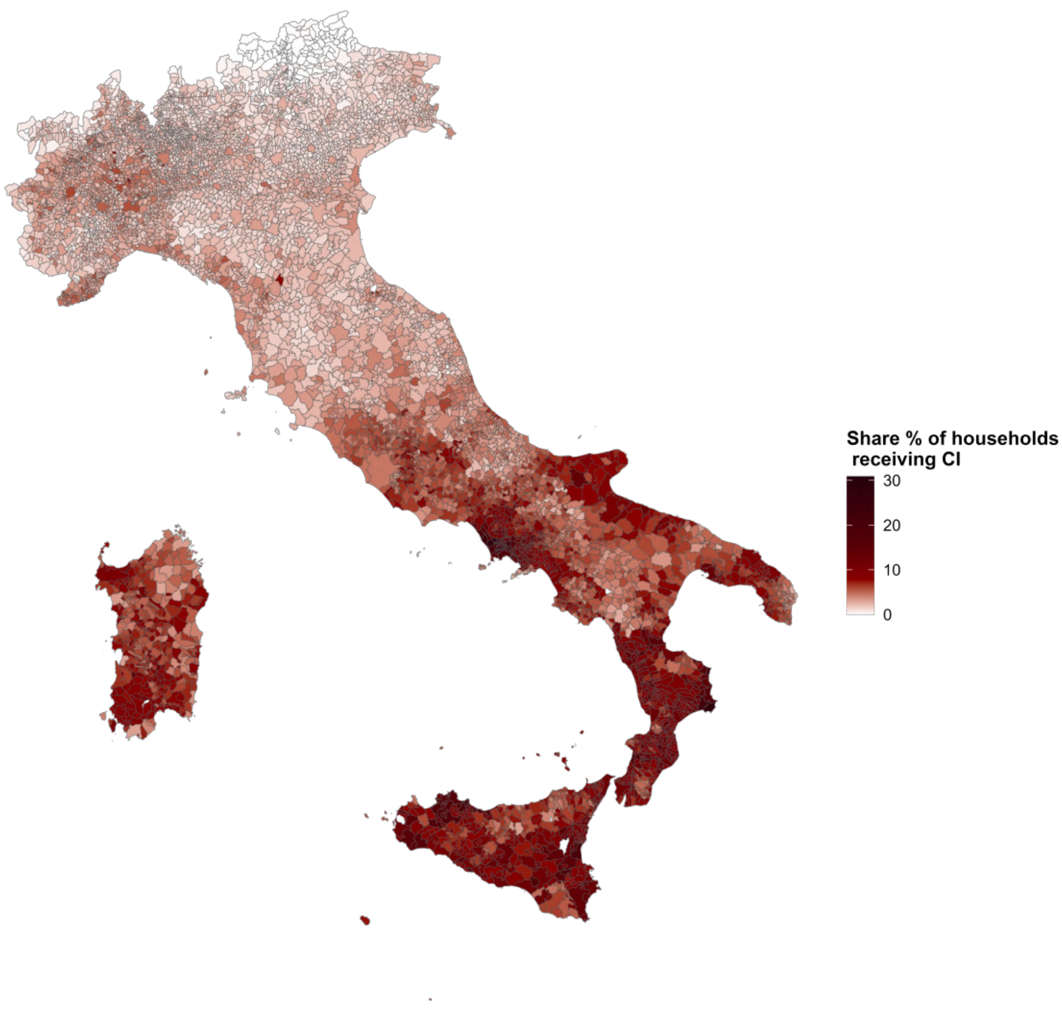


(c) 2018-2022



Source: Author's processing on Ministry of the Interior data.

Figure 4.2: Spatial Distribution of Average Municipal Shares of Households receiving the Italian Citizenship Income of Total Households 2019-2021



Source: Author's processing of INPS data.

4.3 Empirical Strategy and Data

4.3.1 Empirical Strategy

To estimate the causal relationship between Italian Citizenship Income and the share of votes received by the incumbent party, we exploit the municipal heterogeneity in the share of households receiving the transfer (observable in Figure 4.2) to set up a Difference in Difference (DiD) study design. Following the seminal approach of Card (1992) and many econometric works on the study of the causal parameters, such as Cerulli (2015), Filippetti and Cerulli (2018), Callaway and Huang (2020), Cerulli and Ventura (2021), D'Haultfœuille et al. (2023), Callaway et al. (2024), we implement a DiD with continuous treatment, where the treatment intensity variable is, therefore, represented by the share of families receiving the IC.

The basic approach of DiD is to compare the difference in outcomes after and before the intervention between two groups (treated and untreated). The use of DiD allows us to avoid endogeneity problems typical of comparisons between heterogeneous groups (Meyer, 1995). This assumption can be formally expressed as:

$$\Delta Y = [E(Y_{post} \mid D = 1) - E(Y_{pre} \mid D = 1)] - [E(Y_{post} \mid D = 0) - E(Y_{pre} \mid D = 0)] \tag{4.1}$$

where:

- D is a dummy variable indicating whether a unit was treated ($D = 1$) or untreated ($D = 0$);
- Y_{pre} represents the value of the variable of interest in the pre-treatment period;
- Y_{post} represents the value of the variable of interest in the post-treatment period;
- $E(Y_{post} \mid D = 1)$ is the expectation of the variable of interest for the group treated after the intervention;
- $E(Y_{pre} \mid D = 1)$ is the expectation of the variable of interest for the group treated before the intervention;
- $E(Y_{post} \mid D = 0)$ is the expectation of the variable of interest for the control group after the intervention;
- $E(Y_{pre} \mid D = 0)$ is the expectation of the variable of interest for the control group before the intervention.

When we are implementing a DiD study design with continuous treatment, the treatment is modelled as a continuous variable, and every unit is identified by the intensity of the policy (Card, 1992; Callaway and Huang, 2020; Callaway et al., 2024), which in our case is the share of households treated by the CI. This assumption can be formally expressed as:

$$\Delta Y = [E(Y_{post} \mid D) - E(Y_{pre} \mid D)] - [E(Y_{post} \mid D = 0) - E(Y_{pre} \mid D = 0)] \tag{4.2}$$

where:

- D is a continuous variable that represents the intensity of the treatment, in this case the share of households receiving italin income support;
- Y_{pre} represents the value of the variable of interest in the pre-treatment period, in this case the share of votes received by the incumbent party;
- Y_{post} represents the value of the variable of interest in the post-treatment period, in this case the share of votes received by the incumbent party;
- $E(Y_{post} \mid D)$ is the expectation of the variable of interest for the group treated after the intervention;
- $E(Y_{pre} \mid D)$ is the expectation of the variable of interest for the group treated before the intervention;
- $E(Y_{post} \mid D = 0)$ is the expectation of the variable of interest for the control group after the intervention;
- $E(Y_{pre} \mid D = 0)$ is the expectation of the variable of interest for the control group before the intervention.

Formally, we have a population of municipalities observed over three years i.e. 2013, 2018, and 2022. In 2013, the IP presented itself for the first time in the national political elections, whereas in 2018, the proposing party obtained the most seats in Parliament as a single party and the CI was introduced in 2019. Consequently, 2013 and 2018 represent in our estimates the pre-treatment years. The year of Italy's first general election after the introduction of CI (post-treatment) is 2022. Furthermore, to exclude potential anticipation effects on citizens' electoral behaviour towards politics, we use 2008 as an additional pre-treatment electoral year treatment. This election year precedes the first electoral participation of the IP in the parliamentary renewal elections and for this reason, for these latest estimates, we consider the votes obtained by the other "not incumbent parties" (hereinafter also Not IP).

In our study estimated model is:

$$ShareVotesIP_{it} = \theta_i + \lambda_t + \beta_0 X_{it} + \beta_1 CI_{ik} + \varepsilon_{it} \tag{4.3}$$

where:

- $ShareVotesIP_{it}$ is the outcome of interest in municipality i at year t. We test the effect of policy on the share of votes taken by the IP;
- θ_i is municipal-level fixed effect;
- λ_t is time-level fixed effect;
- X_{it} is a time-varying vector of municipal-level covariates, including economic, socio-demographic, environmental and institutional variables (see Table 4.3 for descriptive statistics). Specifically, the covariates are: *Average municipal per capita income* (average amount of municipal income declared over total taxpaying citizens); *Municipal poverty share* (citizens declaring less than 10 thousand euros in income of total taxpaying citizens); *Share of municipal rich* (citizens declaring more than 120 thousand of total taxpaying citizens); *Resident population* (inhabitant living in Italian municipalities as of December 31 of each year); *Share of immigrants* (number of immigrants out of total resident citizens); *Share of the population over 18 years old* (citizens voters); *Land consumption per capita* (land used out of total resident population); *Population density* (resident population on total surface municipal area); *Number of plants of enterprises*

(local units of enterprises); *Average number of employees of the plant* (workers average of local units of enterprises); *Female mayor* (dummy variables. If the mayor is female it is = 1, otherwise it is = 0); *Graduate mayor* (dummy variables. If the mayor has a degree it is = 1, otherwise it is = 0); *Graduated municipal councillors* (dummy variables. If the majority of councillors are graduated it is = 1, otherwise it is = 0); *Age of mayor* (average age of the Mayor); *Average age of councillors* (average age of municipal councillors); *Average age of city council* (average age of municipal asessors); *Mayors under 40* (dummy variables. If the mayor is under 40 years old it is = 1, otherwise it is= 0); *Mayors under 50* (dummy variables. If the mayor is over 50 years old it is = 1, otherwise it is = 0); *Civic lists* (dummy variables. If the administering list is civic it is = 1, otherwise it is = 0); *City Council of IP* (dummy variables. If the list administering is from the IP it is = 1, otherwise it is = 0).

- CI_{ik} is the share of households receiving the income support measure in municipality i at time k. In this case, k represents the time between 2018 and 2022. As the policy was launched in 2019, we take the average of 2019, 2020, and 2021. As a robustness check, we also test each year separately;
- ε_{it} is a error.

The specification 39 is the generalised DiD with continuous treatment (Card, 1992; Callaway and Huang, 2020; Callaway et al., 2024). The parameter of interest is β_1, which captures the effect of the continuous treatment (i.e. share of families receiving the CI) on the share of the vote taken by the IP at the municipal level. To strengthen the plausibility of the exogeneity assumption, since unobservable local-level variables and time-common trends may bias the coefficients of the variables, our model accounts for municipal and year-fixed effects (Hansen, 2007; De Chaisemartin and d'Haultfoeuille, 2020; Imai and Kim, 2021).

Practically, our econometric estimates focus on studying the influence of the CI on the electoral results of the IP in Italian municipalities, considering both the votes obtained in the Chamber of Deputies and for the Senate, in order to obtain complete estimates of voter behaviour in Italy[8] (see Section 4.4).

To test whether socio-spatial and institutional heterogeneity influences the electoral impact of the Italian Citizenship Income on votes for the IP (**Research question RQ**$_2$), we conducted further analyses (presented in detail in Section 4.5). In particular, we have divided the Italian municipalities into two regional macro-areas[9], the Centre-North

[8]In this work the estimates focus on both Italian elective chambers, since the 2018 and 2022 political elections were held with an electoral law different from the one in force in the previous political elections, and since, in the three-year observation period, the electorate and parliamentary representation in both houses have changed significantly, due to institutional reforms. In particular, Electoral Law n. 270 of 21 December 2005, in force in the 2008 and 2013 elections, was an electoral law that provided for a proportional system with a coalition majority bonus, assigned without a minimum threshold of votes. Otherwise, Law 3 November 2017, n. 165 is a mixed electoral system since 37 percent of the seats in each elective chamber are assigned through a majority system, while the remaining 63 percent is distributed proportionally.

[9]The Centre-North with Emilia Romagna, Friuli Venezia Giulia, Liguria, Lombardy, Marche, Piedmont, Sicily, Tuscany, Trentino Alto Adige, Umbria, Val d'Aosta, Veneto. The Centre-South

(aggregating the regions of Northern and Central Italy in the same category) and the South (with the regions of Central-South and the Islands).

We then further divided the municipal units between the municipalities that are above or below the median value of specific economic, social and institutional indices and variables, such as social capital, institutional quality index (IQI), the share of immigrants, Unemployment rate, Social and material vulnerability, Incidence of illiterates, since, following the extensive literature on the topic described in the previous sections, for example Bloise et al. (2021), Faggian et al. (2021), Bloise et al. (2023), they appear, especially in Italy, to be key factors capable of influencing electoral results. We carried out these subdivisions both for the entire sample of Italian municipalities and for the Northern and Southern Regions, to study in depth the effects of heterogeneity.

Furthermore, as anticipated, we produced further analyses including the year of the national political elections preceding the accession of the IP, i.e. 2008, focusing attention on the votes obtained by Not IP (illustrated in Section 4.6.1). The objective of these analyses is, therefore, to exclude possible anticipatory effects of citizens' electoral behaviour regarding the Italian Citizenship Income. Also, in this case, the influence of spatial heterogeneity was evaluated by dividing the municipal data set between North and South.

As robustness analysis, we make estimates using the ratio between IP votes and the total number of voters, as well as the total resident population in each municipality (see Section 4.6.2). We also test the impact of the shares of households receiving CI in each treatment year considered, 2019, 2020, and 2021, to reveal any differences between treatment years. We also aligned the Not IP series estimates with the IP estimates using only the period from the voting year 2013 to the year 2022. Finally, for both the incumbent party and the Not IP, we individually compared the voting years before the introduction of the CI with the year of treatment 2022.

4.3.2 Data and Descriptive Statistics

Table 4.3 reports the descriptive statistics of the variables used in the econometric analyses, relating to the years in which the general political elections for the renewal of the members of the Chamber of Deputies and the Senate of the Republic were held (2008, 2013, 2018 and 2022). In particular, the table presents statistics divided into:

1. 'Outcome variable', with the average incumbent party vote percentages in Italian municipalities out of the total number of voters and for robustness analysis with respect to the resident population and voters only;
2. 'Independent variable' (or treatment variable), with the average municipal percentage shares of families who received CI in the three-year period 2019-2021, out of the total

with Abruzzo, Basilicata, Calabria, Campania, Lazio, Molise, Puglia, Sardinia.

families, and for robustness analysis with the percentage shares of families for the years 2019, 2020 and 2021;
3. ‘Control variables’, economic, social, environmental and institutional covariates;
4. Socio-economic and institutional variables, are used for simple splits.

The descriptive statistics in Table 4.3 reveal the impact of the IP in 2013 (the first electoral round in which it stood for election)[10], the significant growth of its consensus in 2018, as well as the subsequent reduction in votes in 2022. This reduction can be found for both legislative assemblies and the total number of voters entitled to vote and with respect only to voting citizens and the total resident population in each Italian municipality.

The share of votes obtained by the Not IP compared to the total number of voters, therefore excluding the IP, in 2008 (the voting year before the accession of the IP) is greater than 80 percent, in both elective chambers. In this sense, the statistics obtained by comparing other elections make clear the significant impact of the incumbent party on the Italian political landscape. In fact, the presence of the IP in the political elections has significantly reduced the share of votes attributed to all the other parties, as in 2013 their share of votes fell to less than 60 percent and in 2018 to just over 50 percent. Another interesting data refers to the election year in which the CI was present (treatment year). Indeed, in 2022, despite the high loss of total support for the IP, the share of votes obtained by the other parties was similar to that recorded in 2013. This could reveal citizens’ disaffection towards politics, with the consequent growth in abstention.

The statistics related to the independent or treatment variable, which corresponds to the average percentage of households that received the monetary transfer between 2019 and 2021 out of the total number of households, indicate the complete absence of beneficiary households in the years 2013 and 2018, as the support measure only commenced the processing of the first beneficiaries in 2019. From this point of view, as the data show, between 2019 and 2021, a significant and growing number of Italian households took advantage of this benefit.

The growth of beneficiaries is detailed by the annual specifications relating to the number of beneficiary families. The highest percentage of families treated by the CI was reached at the municipal level in 2021 (the average share of recipients reached 6 percent of the total families). In particular, in some municipalities in Southern Italy, the number of families treated exceeded half of the total families.

Regarding the other municipal variables considered as controls, there are significant changes between the reference years. There is an increase in the average municipal income

[10] The party proposing the Italian Citizenship Income was founded in 2009 and presented himself for the first time in the general elections for the renewal of both houses of the Italian Parliament in 2013.

per capita of more than 730 Euros between 2018 and 2022 and almost 1800 Euros across 2018-2022. The standard deviation of incomes between the years of analysis also increased. There is also a decrease in the growth of social deprivation, as the share of citizens in poverty (with an income of less than 10 thousand Euros) increased by approximately 1.7 percent between 2018 and 2022 and by more than 4 percent between 2013-2022. A clear decrease (more than 8 percent) was recorded compared to 2008. The share of citizens who declare wealth over 120 thousand Euros also increased.

The demographic variables, resident population, the share of citizens over 18 years of age, and the share of immigrants, remained substantially stable in the years considered. Statistics show an increase in employment and work. For example, the share of enterprises and the average percentage of companies' employees increased significantly. The land consumed per inhabitant also increased significantly, by almost one and a half hectares between 2008 and 2022.

In terms of institutional variables, there is a generational change in the leadership of city halls in the years considered. The number of female mayors and graduate mayors has increased in the years considered. The average age of mayors has decreased from 63.67 to 55.96 years in less than ten years. The percentage of winning civic lists has remained stable, indicating a continuity in civic participation in municipal elections. This has also affected the age of representation within council assemblies, which is very small in the years of observation.

Finally, Table 4.3 shows the median values of socioeconomic variables used to carry out simple splits to estimate the influence of the heterogeneity of certain key factors on the impact of the Italian Citizenship Income on the consensus of the ruling party. In particular, these variables are 1) the synthetic Institutional Quality Index (IQI)[11], created by Nifo and Vecchione (2014); 2) the level of social capital on a provincial scale[12]; 3) the unemployment rate at the provincial level; 4) share of immigrants at municipal level; 5) the incidence of municipal illiterates[13]; 6) the municipal social and material vulnerability[14].

[11] IQI is calculated using composite sub-indicators, referring to the provincial level of corruption, government effectiveness, quality of regulation, rule of law, voice and responsibility. The use of this indicator is widely validated at the level of scientific, economic and statistical literature, for example in the studies conducted by Ferrara and Nisticò (2019), D'Ingiullo and Evangelista (2020), Alfano and Ercolano (2021), De Luca et al. (2021), Corradini (2021) and Mosconi and D'Ingiullo (2023).

[12] The social capital is represented by the weight of cooperative companies, or by the number of employees of cooperative companies out of the total number of employees at provincial level.

[13] The incidence of illiteracy at the municipal level, elaborated by ISTAT starting from census data, refers to the latest available year, 2011. It is important to highlight that this year represents the latest data available at the time of drafting this document. Accessible at the link: https://ottomilacensus.istat.it/.

[14] The social and material vulnerability index at the municipal level, developed by ISTAT starting from census data, refers to the latest available year, 2011 (the latest data available at the time of drafting this document). Accessible at the link: https://ottomilacensus.istat.it.

Table 4.3: Descriptive Statistics on Variables in the National General Election Years 2008-2022

Variables	***Pre-Treatment***			***Post-Treatment***
	2008	**2013**	**2018**	**2022**
Outcome variable				
Votes IP Chamber of Deputies out of total voters	-	16.97(5.46)	19.94(7.25)	7.2(4.61)
Votes IP Senate out of total voters	-	15.69(5.27)	19.71(7.07)	7.2(4.55)
- For robustness				
Votes Not IP Chamber out of Deputies on total voters	81.14(6.32)	57.9(7.82)	53.56(10.81)	57.14(11.16)
Votes Not IP Senate out of total voters	80.97(6.38)	58.91(7.85)	53.66(10.76)	57.14(11.13)
Votes IP Chamber out of Deputies of total voters	-	22.65(6.87)	27.54(10.92)	11.86(8.66)
Votes IP Senate out of total votes	-	21(6.62)	27.29(10.71)	11.86(8.56)
Votes IP Chamber of Deputies out of total population	-	13.52(4.33)	16.18(6.22)	5.86(3.86)
Votes IP Chamber of Deputies out of total population	-	11.53(3.93)	14.76(5.54)	5.86(3.81)
Treatment variable				
Average share of households receiving CI of total families 2019-2021	0(0)	0(0)	0(0)	4.04(3.46)
- For robustness				
Share of households receiving CI of total families 2019	0(0)	0(0)	0(0)	2.61(2.22)
Share of households receiving CI of total families 2020	0(0)	0(0)	0(0)	3.5(3.07)
Share of households receiving CI of total families 2021	0(0)	0(0)	0(0)	5.99(5.34)
Control variables				
Economic				
Average municipal per capita income	15447.81(3477.83)	16489.02(3531.99)	17530.19(3748.31)	18279.8(3773.4)
Municipal poverty share	38.27(12.49)	33.89(10.87)	31.43(10.09)	29.74(9.25)
Share of municipal rich	0.35(0.041)	0.37(0.42)	0.48(0.5)	0.58(0.53)
Resident population	7556.27(40656.79)	7676.34(42279.2)	7581.84(42866.51)	7486.28(41801.1)
Share of immigrants	4.94(3.62)	6.45(4.48)	6.57(4.19)	6.7(4.42)
Geo-Demographic				
Share of population over 18 years old	83.55(4.05)	84.55(3.68)	85.65(3.81)	85.91(3.65)
Land consumption per capita	6.37(4.37)	6.65(4.86)	7.06(5.39)	7.36(5.8)
Population density	30247.62(64316.8)	30764.77(65073.96)	30413.84(64678.74)	30099.74(64082.54)
Industrial				
Number of plants of enterprises	-	610.62(4281.87)	604.32(4449.22)	626.23(4595.33)
Average of number of employees of the plants	-	2092.11(15707.39)	2196.74(17732.72)	2237.98(17739.01)
Governance				
Female mayor	0.1(0.3)	0.17(0.37)	0.25(0.43)	0.27(0.44)
Graduate mayor	0.42(0.49)	0.49(0.5)	0.54(0.5)	0.54(0.5)
Graduated municipal councilors	0.05(0.22)	0.08(0.27)	0.1(0.3)	0.1(0.3)
Age of mayor	67.66(9.48)	63.67(9.6)	58.89(9.6)	55.96(9.93)
Average age of councillors	62.65(4.27)	57.84(4.6)	53(4.8)	50.2(4.88)
Average age of municipal asessors	63.61(7.1)	58.7(7.79)	53.81(9.81)	51.22(9.6)
Mayors under 40	0.0003(0.02)	0.02(0.13)	0.08(0.28)	0.15(0.35)
Mayors over 50	0.96(0.19)	0.91(0.28)	0.85(0.36)	0.77(0.42)
Civic lists	0.86(0.35)	0.02(0.14)	0.96(0.2)	0.97(0.17)
City Council of IP	-	0.01(0.1)	0.04(0.2)	0.01(0.01)
- For simple split				
IQI	0.68	0.68	0.69	0.69
Social Capital	3.86	3.88	3.86	3.67
Unemployment rate	8.38	8.38	8.38	6.6
Share of immigrants	4.24	5.61	5.86	5.91
Incidence of illiterates (values referring to 2011)	0.7	0.7	0.7	0.7
Social and material vulnerability (values referring to 2011)	98.6	98.6	98.6	98.6

Note: Values are reported as mean (standard deviation). For example, "16.97(5.46)" indicates a mean of 16.97 with a standard deviation of 5.46. The statistics relating to the socio-economic and institutional indicators for the sample split represent the median values, calculated taking into account the availability of data in the election years considered. Incidence of illiterates and Social and material vulnerability refer to 2011 census data.

4.4 Results

4.4.1 The Effects on IP Votes in Italian Municipalities

Table 4.4 reports the results of the DID estimates, using the share of votes received by the IP as the dependent variable (i.e. Eq. 39). In all model specifications (Models 1, 2, 3 4), with and without the inclusion of socioeconomic control variables, the coefficient of interest is positive and significant. Consequently, the share of families who received the CI between 2019 and 2021 had a positive impact on the votes for the party that proposed the measure. In other words, the increase in the number of families receiving the CI is associated with a boost in the electoral results for the IP.

Specifically, the models without controls (Model 1 for the Chamber of Deputies and Senate in Table 4.4) present the highest coefficients, relative to the treatment variable, since the coefficients of the regressions with controls present show significantly lower values and closer to zero (in particularly in Models 3 and 4 for both elected chambers). For example, the economic variables (in our study the average per capita income, the municipal poverty rate and the share of rich people) reduce the influence of the CI on votes (see Model 2, both for the Chamber of Deputies and for the Senate). The addition of other socio-demographic, industrial controls, such as production sites and their workers, and local government variables, such as female mayors, age of city councillors, etc., help to further reduce the magnitude of the positive impact of CI on the ruling party consensus (Models 3 and 4).

In general, the difference in magnitude of the coefficients of Eq. 39, between the different specifications of the regression model, highlights the relative impacts that the socioeconomic variables have in influencing the voting behaviour of citizens considering the treatment of the CI, as shown by Bloise et al. (2021), Faggian et al. (2021), Bloise et al. (2023), Giuliani (2023).

Specifically, our results demonstrate that the CI can influence electoral support for the incumbent party in general elections (**Research Question $\mathbf{RQ}_1$**), even if the magnitude of the association is influenced by socioeconomic variables.

Finally, municipalities with a higher percentage of beneficiaries are associated with positive support for the IP. This is in line with the theory that identifies voters as economic agents seeking to maximise their benefits, such as Lewis-Beck (1985), Manacorda et al. (2011), Elinder et al. (2015). The impact of CI in diverse socio-economic contexts leads to the exploration of our second research question (**Research question $\mathbf{RQ}_2$**) in the next Section, which delves into the influence of socioeconomic and institutional heterogeneity.

Table 4.4: Regression Results on the Total Italian Municipalities

	Chamber of Deputies				Senate			
	Model 1	**Model 2**	**Model 3**	**Model 4**	**Model 1**	**Model 2**	**Model 3**	**Model 4**
Average share of households receiving CI of total families 2019-2021	0.437***	0.307***	0.268***	0.269***	0.458***	0.332***	0.292***	0.293***
	(0.017)	(0.02)	(0.02)	(0.021)	(0.017)	(0.02)	(0.019)	(0.02)
Eco		✓	✓	✓		✓	✓	✓
Eco/Geo-Demo			✓	✓			✓	✓
Eco/Geo-Demo/Ind/Gov				✓				✓
Municipal Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	23815	17561	17406	16884	23815	17561	17406	16884

Note: All continuous predictors are mean-centred and scaled by 1 standard deviation. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

4.5 Heterogeneity of the Results

4.5.1 Macroregional Heterogeneity

In this section, we perform a sample split dividing municipalities between Centre-North and Southern regions. Municipalities in Northern regions show opposite trends to those obtained considering all the Italian municipalities (see Table 4.5): the coefficients of interest in Eq. 39 are negative and significant across all models. This result suggests that, in more economically prosperous municipalities where fewer households are benefiting from CI, the electorate tends not to reward the political force that promoted the measure (Giuliani, 2023; Albanese et al., 2024).

The regressions without controls present a very negatively marked magnitude of the coefficients in absolute value (Models 1), which reveals a direct negative association between the treatment of the CI and the votes for the incumbent party. Even the coefficients of the independent variable relating to the specifications that present together economic, socio-demographic, industrial, and local government controls (Models 3 and 4 for both axes), present very similar values in terms of intensity, revealing a limited influence of these variables on the CI contribution to the votes for the ruling party in the municipalities of Northern Italy. However, the coefficients of Models 2, which present only economic variables as controls, we can note, even in the Northern regions, a high influence of these variables on the impact of CI on votes. Additionally, the magnitude of the treatment variable is markedly lower - in absolute value - compared to the coefficients obtained in the other specifications of the model (see Models 2 for both the Chamber of Deputies and the Senate).

However, if we consider only the municipalities of Southern Italy, the coefficients of the independent variable in Eq. 39 are positive and statistically significant in all the estimated models, as in the regressions conducted on the total of Italian municipalities (see Table 4.5). In other words, the increase in the average share of families benefiting from the CI programme is positively correlated with the increase in votes for the IP.

In the South, unlike in the North, the control variables appear to be key drivers for evaluating the impact of the measure on support for the ruling party. In fact, the addition of socio-economic, industrial and institutional controls, albeit to a limited extent (see Models 3 and 4), increases the correlation between the dependent variable of Eq. 39 and the treatment variable. The coefficients from the regressions with only economic controls do not change the association between CI and IP. This last item of evidence is linked to the results obtained in the estimates in the municipalities of the North and on the entire sample of Italian municipalities, since the impact of these income variables in Southern regions is lower than in other areas of the country in consideration of the fact that the distribution of income is more uniform, with less marked disparities and lower incomes.

To some extent, these results are in line with the studies conducted by Albanese et al. (2024), which show greater support in Southern regions for parties promoting welfare subsidies. Furthermore, this evidence confirms the results of works revealing a large consensus for the party that proposed the CI in territories where social exclusion and the demand for assistance and protection from poverty and of the "within inequalities" are higher.

Overall, our results highlight that electoral behaviour diverges between the North and the less developed South of Italy, as in Albanese et al. (2024). In Northern Italy, voter behaviour results more in line with a sociotropic and negatively retrospective economic voting approach (Fiorina, 1978; Levitt and Snyder Jr, 1997). Voters did not reward the incumbent party that created the CI. Voting behaviour seems to be more influenced by overall economic and social conditions, rather than by direct personal monetary transfers. However, Southern Italy presents a more complex dynamic. In fact, in these regions, voting behaviour appears to be influenced by a combination of pocketbook and sociotropic voting (Lewis-Beck, 1985; Lewis-Beck and Stegmaier, 2000, 2008, 2018). It seems that the multidimensional deprivations present in the South, ranging from economic issues to social challenges, make voters more sensitive to the effects of policies on their personal or family income and the general socioeconomic conditions of the community. On the one hand, voting behaviour seems to be influenced by policies that have a direct impact on their well-being (i.e. CI), manifesting a pocketbook-type vote. On the other hand, they are also influenced by a sociotropic vision, evaluating policies in terms of impact on the community. In these respects, the results in the municipalities of the South show a positive retrospective vote, since voters tend to re-vote for the incumbent party that introduced CI.

In short, our analysis shows that spatial heterogeneity and geographical characteristics in Italy are reflected in electoral behaviour, highlighting the complexity of the connection between income support policies and voting (Giuliani, 2023; Albanese et al., 2024).

Regarding the second research question of this study $\mathbf{RQ}_2$, Tables 4.5 and 4.6, show that macro-regional heterogeneity is a decisive driver to explain the impact of the Italian

Citizenship Income on the electoral support of the proposing party. On a general level, considering all Italian municipalities, and in particular, only those in the South, this policy represented a boost in terms of votes for the IP. Conversely, in the North, this monetary transfer led to a loss of electoral consensus.

Table 4.5: Results of the Regression on the Municipalities of Northern Italy

	Chamber of Deputies				Senate			
	Model 1	**Model 2**	**Model 3**	**Model 4**	**Model 1**	**Model 2**	**Model 3**	**Model 4**
Average share of households receiving CI of total families 2019-2021	-0.638***	-0.583***	-0.642***	-0.625***	-0.556***	-0.485***	-0.546***	-0.522***
	(0.046)	(0.054)	(0.055)	(0.055)	(0.045)	(0.053)	(0.054)	(0.055)
Eco		✓	✓	✓		✓	✓	✓
Eco/Geo-Demo			✓	✓			✓	✓
Eco/Geo-Demo/Ind/Gov				✓				✓
Municipal Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	15036	11759	11615	11419	15036	11759	11615	11419

Note: All continuous predictors are mean-centered and scaled by 1 standard deviation. *** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$.

Table 4.6: Results of the Regression on the Municipalities of Southern Italy

	Chamber of Deputies				Senate			
	Model 1	**Model 2**	**Model 3**	**Model 4**	**Model 1**	**Model 2**	**Model 3**	**Model 4**
Average share of households receiving CI of total families 2019-2021	0.459***	0.454***	0.475***	0.486***	0.449***	0.45***	0.469***	0.477***
	(0.026)	(0.033)	(0.033)	(0.036)	(0.025)	(0.032)	(0.032)	(0.034)
Eco		✓	✓	✓		✓	✓	✓
Eco/Geo-Demo			✓	✓			✓	✓
Eco/Geo-Demo/Ind/Gov				✓				✓
Municipal Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	8779	5802	5791	5465	8779	5802	5791	5465

Note: All continuous predictors are mean-centered and scaled by 1 standard deviation. *** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$.

4.5.2 Socioeconomic and Institutional Heterogeneity

The estimates conducted to evaluate the influence of the Italian Citizenship Income on votes for the ruling party, taking into account spatial heterogeneity concerning key social, economic and institutional factors, are presented in Tables 4.7, 4.8, 4.9. They largely confirm the evidence obtained at a macro-regional level between North and South (in Tables 4.4, 4.5 and 4.6). These analyses allow us to delve into the disparities present in Italy in greater detail since the country is divided, not only in geographical terms, but also concerning key factors of territorial and local development (such as, for example, the level of corruption synthesised in IQI, or the unemployment rate).

At a national level, in the municipalities of the provinces below the median value of the Institutional Quality Index (IQI) (Nifo and Vecchione, 2014), the coefficients of the treatment variable of Eq. 39, both for the Chamber of Deputies and for the Senate, demonstrate that the Italian Citizenship Income is positively associated with the electoral outcome of the governing party. Contrastingly, in territories with higher institutional quality (above the median), there is a negative association between Italian citizenship income and electoral support for the IP.

Similar results are obtained with the estimates conducted on Italian municipalities that have an unemployment rate (calculated at the provincial level) higher than the national median level. In these municipalities, the CI has determined a positive and significant direct consensus for the IP, both for the Chamber of Deputies and for the Senate (Model 1). By contrast, negative and significant results are found in municipalities with a lower provincial unemployment rate (below the general median level).

The regressions in Table 4.7 show, both for the Chamber of Deputies and for the Senate, that the CI determined an increase in consensus for the party that proposed the measure, in the municipalities characterised by a greater incidence of illiteracy among citizens and a higher rate of social and material vulnerability (higher than the national median level). In contrast, in municipalities characterised by a lower presence of illiterate citizens and a lower level of social vulnerability, negative coefficients are recorded about the magnitudes of the coefficients of the treatment variable. Furthermore, the coefficients below the median of the social and material vulnerability index are weakly significant, with a magnitude close to zero (Model 1). This highlights the crucial role of social and material vulnerability as a determining factor in influencing the effect of CI on support for the proposing party.

The results obtained by dividing the sample by level of Social Capital and Share of immigrants do not highlight any heterogeneity of the results: positive and significant coefficients persist for both the Chamber of Deputies and the Senate.

The evidence in Table 4.7 show a strong influence of the heterogeneity of economic, social and institutional factors in Italian Municipalities (in particular regarding the quality of institutions, the unemployment rate, the incidence of illiterate citizens and the level of social and material vulnerability), as factors capable of significantly impacting on the electoral consensus in Italy, as demonstrated by Bloise et al. (2021), Faggian et al. (2021), Bloise et al. (2023) and Giuliani (2023).

Less heterogeneous results, in terms of the sign of the coefficients, are obtained by dividing the sample with respect to the median values of the social, economic and institutional indicators within the Italian geographical areas (North and South).

Considering the municipalities of Northern Italy (see Table 4.8), in almost all models (above and below the median) and for both elective chambers, negative and significant values are observed regarding the influence of the CI on votes for the incumbent party. However, although the regressions shown in Table 4.5 largely highlight the absence of a positive impact of the treatment variable of Eq. 39 on the IP consensus, focusing on the

Table 4.7: Results of Regressions Above and Below of Median of Scio-Economic and Institutional Factors (Italy)

	IQI		Social Capital		Unemployment rate		Share of immigrants		Incidence of illiterates		Social and material vulnerability	
Chamber of Deputies												
	< median	**< median**	**< median**	**< median**	**< median**	**< median**	**< median**	**< median**	**< median**	**< median**	**< median**	**< median**
	Model 1	**Model 2**	**Model 1**	**Model 2**	**Model 1**	**Model 2**	**Model 1**	**Model 2**	**Model 1**	**Model 2**	**Model 1**	**Model 2**
Average share of households receiving CI of total families 2019-2021	0.482***	-0.502***	0.435***	0.157***	-0.535***	0.507***	0.277***	0.107***	-0.345***	0.407***	-0.123**	0.332***
	(0.033)	(0.068)	(0.032)	(0.03)	(0.068)	(0.032)	(0.032)	(0.03)	(0.047)	(0.031)	(0.052)	(0.029)
Eco/Geo-Demo/Ind/Gov	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Senate												
	IQI		**Social Capital**		**Unemployment rate**		**Share of immigrants**		**Incidence of illiterates**		**Social and material vulnerability**	
	< median	**< median**	**< median**	**< median**	**< median**	**< median**	**< median**	**< median**	**< median**	**< median**	**< median**	**< median**
	Model 1	**Model 2**	**Model 1**	**Model 2**	**Model 1**	**Model 2**	**Model 1**	**Model 2**	**Model 1**	**Model 2**	**Model 1**	**Model 2**
Average share of households receiving CI of total families 2019-2021	0.482***	-0.367***	0.474***	0.16***	-0.409***	0.502***	0.306***	0.142***	-0.256***	0.414***	-0.049	0.333***
	(0.032)	(0.066)	(0.031)	(0.029)	(0.067)	(0.032)	(0.031)	(0.029)	(0.046)	(0.03)	(0.051)	(0.029)
Eco/Geo-Demo/Ind/Gov	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Municipal Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	7286	9229	8930	7650	9198	7317	7644	9067	9877	6834	9098	7613

Note: All continuous predictors are mean-centred and scaled by 1 standard deviation. *** p < 0.01; ** p < 0.05; * p < 0.1.

analysis of the intensity and magnitude of the coefficients, different levels of heterogeneity can still be observed.

Evidence of socioeconomic heterogeneity can be observed from the coefficients of the regressions conducted on the institutional quality indicator, both for the Chamber of Deputies and for the Senate. The coefficients of the municipalities that present a level of institutional quality lower than the median level are strongly negative and significant (Models 1 for the IQI). Above the median level, they are, instead, positive and significant (Models 2). This trend may depend on a positive inclination of citizens of Northern municipalities that have a higher institutional quality towards economic policies aimed at protecting vulnerable and deprived people.

As regards the share of immigrants (in both elective seats), in the municipalities of the North that are above the median level of this indicator, non-statistically significant coefficients are recorded, with intensities close to zero, especially for the Senate. In contrast, municipalities with a value lower than the median level record negative coefficients (with approximately one point of difference compared to municipalities above the median of the immigrant share). These differences highlight how the migration issue represents a significant driver in conditioning the impact of the CI on votes for the proposing party, especially in territorial contexts characterised by better levels of wellbeing and economic conditions, such as the regions of Northern Italy.

Large levels of heterogeneity are also detected in the coefficients obtained considering the median level of provincial social capital. In fact, in municipalities that have a social capital value below the median level, coefficients with a magnitude close to zero are recorded (see Model 1 for the Chamber of Deputies and Senate).

Similarly, it is possible to detect heterogeneity in the magnitudes of the coefficients, below and above the median level, in relation to the provincial unemployment rate, the incidence of illiteracy, and the rate of social and material vulnerability on a municipal scale. In municipalities characterised by greater exclusion from the labour market (above the median of the provincial unemployment rate) the coefficients, for the Chamber of Deputies and the Senate, are approximately one and a half points lower than the coefficients of municipalities below the median level. This shows that the lack of employment conditions impacts on consensus less negatively.

In municipal contexts with vulnerability rates higher than the median value, a negative and significant correlation is recorded, with a marked intensity between the income support measure and consensus for the ruling party. This may depend on unobservable factors, such as the presence, in these municipalities, of other forms of support or structured protection systems, which reduce the potential link between politics and votes.

Therefore, in light of these results, there seems to be evidence that, even in the municipalities of Northern Italy alone, the heterogeneity of socioeconomic and institutional factors has a significant influence on the effect of the Italian Citizenship Income on IP consensus, answering positively to **Research question $\mathbf{RQ}_2$**.

In the municipalities of the South, the heterogeneity of the social, economic and institutional variables only has a small influence on the coefficients of the independent variable (see Table 4.9), since the latter are positive and significant in all the regressions carried out on the indicators considered. Unlike that of the Northern municipalities, the magnitude of the estimates for the Southern municipalities is very similar, both above and below the median values. This result may be partially attributed to the unfavourable economic and social conditions that characterise the territories of the South, which have lower economic performances, lower efficiency of the institutions as well as a greater dependence on forms of state subsidies, compared to the regions of the North, as highlighted by Albanese et al. (2024).

In general, these results highlight how the different territorial realities and certain key factors significantly influence the behaviour of voters. In territories characterised by greater social, economic, and institutional deficiencies, votes are more linked to individual or family direct monetary transfers and subsidies. Overall, these results confirm that there is a strong role of socioeconomic and institutional heterogeneity (i.e. **Research Question $\mathbf{RQ}_2$**).

Table 4.8: Results of Regressions Above and Below of Median of Socio-Economic and Institutional Factors (North)

	Chamber of Deputies											
	IQI		**Social Capital**		**Unemployment rate**		**Share of immigrants**		**Incidence of illiterates**		**Social and material vulnerability**	
	< median	**< median**	**< median**	**< median**	**< median**	**< median**	**< median**	**< median**	**< median**	**< median**	**< median**	**< median**
	Model 1	**Model 2**	**Model 1**	**Model 2**	**Model 1**	**Model 2**	**Model 1**	**Model 2**	**Model 1**	**Model 2**	**Model 1**	**Model 2**
Average share of households receiving CI of total families 2019-2021	-0.962***	0.362***	-0.243***	-0.856***	-1.08***	-0.537***	-1.097***	-0.093	-0.776***	-0.33***	-0.298***	-0.946***
	(0.075)	(0.112)	(0.078)	(0.084)	(0.115)	(0.089)	(0.097)	(0.072)	(0.088)	(0.076)	(0.096)	(0.072)
Eco/Geo-Demo/Ind/Gov	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
	Senate											
	IQI		**Social Capital**		**Unemployment rate**		**Share of immigrants**		**Incidence of illiterates**		**Social and material vulnerability**	
	< median	**< median**	**< median**	**< median**	**< median**	**< median**	**< median**	**< median**	**< median**	**< median**	**< median**	**< median**
	Model 1	**Model 2**	**Model 1**	**Model 2**	**Model 1**	**Model 2**	**Model 1**	**Model 2**	**Model 1**	**Model 2**	**Model 1**	**Model 2**
Average share of households receiving CI of total families 2019-2021	-0.876***	0.52***	-0.08	-0.776***	-0.931***	-0.469***	-0.974***	-0.005	-0.65***	-0.266***	-0.186*	-0.886***
	(0.073)	(0.111)	(0.078)	(0.08)	(0.112)	(0.088)	(0.094)	(0.071)	(0.086)	(0.075)	(0.095)	(0.07)
Eco/Geo-Demo/Ind/Gov	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Municipal Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	5266	5994	5910	5350	5976	5284	5379	5881	5642	5618	5786	5474

Note: All continuous predictors are mean-centered and scaled by 1 standard deviation. *** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$.

Table 4.9: Results of Regressions Above and Below of Median of Socio-Economic and Institutional Factors (South)

	Chamber of Deputies											
	IQI		**Social Capital**		**Unemployment rate**		**Share of immigrants**		**Incidence of illiterates**		**Social and material vulnerability**	
	< median	**< median**	**< median**	**< median**	**< median**	**< median**	**< median**	**< median**	**< median**	**< median**	**< median**	**< median**
	Model 1	**Model 2**	**Model 1**	**Model 2**	**Model 1**	**Model 2**	**Model 1**	**Model 2**	**Model 1**	**Model 2**	**Model 1**	**Model 2**
Average share of households receiving CI of total families 2019-2021	0.339***	0.574***	0.576***	0.263***	0.718***	0.205***	0.413***	0.441***	0.406***	0.5***	0.596***	0.312***
	(0.05)	(0.089)	(0.047)	(0.057)	(0.107)	(0.065)	(0.056)	(0.049)	(0.047)	(0.059)	(0.079)	(0.046)
Eco/Geo-Demo/Ind/Gov	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
	Senate											
	IQI		**Social Capital**		**Unemployment rate**		**Share of immigrants**		**Incidence of illiterates**		**Social and material vulnerability**	
	< median	**< median**	**< median**	**< median**	**< median**	**< median**	**< median**	**< median**	**< median**	**< median**	**< median**	**< median**
	Model 1	**Model 2**	**Model 1**	**Model 2**	**Model 1**	**Model 2**	**Model 1**	**Model 2**	**Model 1**	**Model 2**	**Model 1**	**Model 2**
Average share of households receiving CI of total families 2019-2021	0.35***	0.463***	0.556***	0.27***	0.582***	0.202***	0.416***	0.426***	0.416***	0.462***	0.51***	0.319***
	(0.047)	(0.087)	(0.046)	(0.056)	(0.104)	(0.061)	(0.055)	(0.047)	(0.046)	(0.058)	(0.078)	(0.043)
Eco/Geo-Demo/Ind/Gov	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Municipal Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	2604	2651	2669	2651	2713	2542	2709	2742	2997	2454	2863	2588

Note: All continuous predictors are mean-centered and scaled by 1 standard deviation.*** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$.

4.6 Robustness Checks

4.6.1 Excluding Anticipatory Effects. Estimates on the Other Parties

As described in this section, Tables 4.10, 4.11 and 4.12 report the econometric analysis conducted on the share of municipal votes assigned to parties other than the incumbent party (Not IP), on the total voters in each municipality[15] (see maps in Figures D1), including in the analysis the year of election before the announcement of the CI by the IP (i.e. 2008).

The Eq. 40 reports the estimated model for the study of excluding anticipatory effects:

$$ShareVotesNotIP_{it} = \theta_i + \lambda_t + \beta_0 X_{it} + \beta_1 CI_{ik} + \varepsilon_{it} \tag{4.4}$$

where:

- $ShareVotesNotIP_{it}$: is the share of votes assigned to parties other than the incumbent party (Not IP) in municipality i at year t, including the year election before the introduction of CI (e.g. 2008);
- θ_i: is a municipal-level fixed effect;
- λ_t: is a time-level fixed effect;
- X_{it}: is a time-varying vector of municipal-level covariates, including economic, socio-demographic, environmental, and institutional variables;
- CI_{ik}: is the share of households receiving the income support measure in municipality i at time k;
- ε_{i_t}: is the error term.

The results shown in Tables 4.10, 4.11 and 4.12 reveal coefficients of Eq. 40 opposite in terms of sign compared to those obtained in the estimates carried out on the IP party.

In particular, considering all the Italian municipalities (see Table 4.10), significant negative coefficients of the treatment variable are observed, with very high magnitudes in all the specifications considered (Model 1, 2, 3, 4). In other words, in these municipalities, an increase in the number of families covered by politics in Italy corresponds to a significant decrease in consensus towards non-incumbent parties (connected to the evidence of positive retrospective pocketbook voting in Italy, shown in Table 4.6, for the IP list (Lewis-Beck, 1985; Lewis-Beck and Stegmaier, 2000, 2008, 2018). For these parties, the **Research question RQ$_1$** underlying our work is therefore not confirmed.

Furthermore, the negative independent variable reveals, in each specification of the

[15] The Not IP variable represents the sum of the votes assigned to the parties competing with the party proposing the Italian Citizenship Income, on the total voters in each municipality. The number of blank ballots and null ballots are also considered in this sum, in order to fully evaluate the electoral behaviour of citizens.

model, the presence of anticipation effects on the voting behaviour of citizens, in relation to a possible introduction of an income support measure, such as the CI. The presence of anticipation is further confirmed by the comparison with the estimates conducted in the electoral period from 2013 to 2022 on the ruling party (see Table 4.4) and on the Not IP (robustness estimates in Table 4.1). In fact, estimates that include 2008 show a significantly greater decrease in consensus for Not IP. This suggests that, even before the official introduction of the Italian Citizenship Income, there was a long-term electoral movement penalising other parties. In this sense, voters may have perceived or foreseen the introduction of income support measures and have consequently oriented their vote accordingly based on this perception. This result is in line with the growth of consensus around IP in the years between 2009 (year of foundation) and the 2013 political elections.

From the estimates conducted in the Northern municipalities (Table 4.11) it is possible to detect positive and significant coefficients of the treatment variable in all the model specifications and in both elective chambers. These values show a positive association of the CI on the votes for non-proposing parties, i.e. they note that the policy has electorally disadvantaged the ruling party while benefiting the Not IP. The sign and magnitude of the coefficients of the treatment variable also demonstrate the absence of anticipation effects on votes for Not IP in Northern Italy. In fact, the magnitude of the coefficients is lower in absolute value than those obtained considering the incumbent party and the Not IP themselves in the electoral series which does not consider 2008 in the estimates (see Tables 4.4, 4.1).

Consequently, the better socioeconomic conditions of these municipalities avoided anticipatory effects before the announcement of the CI, resulting in a concrete electoral advantage for the Not IP, confirmed by the evidence of negative and retrospective sociotropic voting reported in Table 4.5 for the IP (Fiorina, 1978; Levitt and Snyder Jr, 1997).

Results opposite to those obtained in the Northern municipalities can be found in the municipalities of Southern regions and the Islands (Table 4.12), where the coefficients of Eq. 39 are negative and significant for the Not IP. In these southern municipalities, there is a negative association between the consensus of non-proposing parties and the CI. Furthermore, at the level compared with the estimates on the entire sample of Italian municipalities, a limited anticipation effect of the consensus regarding the CI is highlighted, given that the magnitude of the coefficients of Eq. 39 is higher than that obtained in the IP and Non-IP estimates, which consider the 2013-2022 series.

These results highlight how in the Southern regions, characterised by a greater pervasiveness of multidimensional phenomena, citizens anticipated the approval of the CI, reducing the consensus attributed to Not IP. Confirm the results on pocketbook and sociotropic voting, retrospectively positive, for the IP in these regions (Albanese et al., 2024), shown in Table 4.6).

In general, these results provide an answer to the second basic question of our work (**Research Question $\mathbf{RQ}_2$**), by demonstrating that territorial heterogeneity has played a decisive role in shaping citizens' electoral behaviour towards income support measures, such as the CI.

Table 4.10: Regression Results on the Total Italian Municipalities on the Not IP

	Chamber of Deputies				**Senate**			
	Rob 1	**Rob 2**	**Rob 3**	**Rob 4**	**Rob 1**	**Rob 2**	**Rob 3**	**Rob 4**
Average share of households receiving CI of total families 2019-2021	-1.012***	-0.905***	-0.883***	-0.882***	-0.998***	-0.9***	-0.879***	-0.877***
	(0.019)	(0.022)	(0.022)	(0.023)	(0.019)	(0.022)	(0.022)	(0.022)
Eco/Geo-Demo/Ind/Gov	✓	✓	✓	✓	✓	✓	✓	✓
Municipal Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	31883	23398	23039	22349	31883	23398	23039	22349

Note: All continuous predictors are mean-centred and scaled by 1 standard deviation. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 4.11: Results of the Regression on the Municipalities of Northern Italy on the Not IP

	Chamber of Deputies				**Senate**			
	Rob 1	**Rob 2**	**Rob 3**	**Rob 4**	**Rob 1**	**Rob 2**	**Rob 3**	**Rob 4**
Average share of households receiving CI of total families 2019-2021	0.411***	0.396***	0.591***	0.259***	0.418***	0.41***	0.584***	0.269***
	(0.058)	(0.067)	(0.074)	(0.068)	(0.058)	(0.066)	(0.073)	(0.068)
Eco/Geo-Demo/Ind/Gov	✓	✓	✓	✓	✓	✓	✓	✓
Municipal Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	20173	15653	15318	15047	20173	15653	15318	15047

Note: All continuous predictors are mean-centered and scaled by 1 standard deviation. *** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$.

Table 4.12: Results of the Regression on the Municipalities of Southern Italy on the Not IP

	Chamber of Deputies				**Senate**			
	Rob 1	**Rob 2**	**Rob 3**	**Rob 4**	**Rob 1**	**Rob 2**	**Rob 3**	**Rob 4**
Average share of households receiving CI of total families 2019-2021	-0.58***	-0.576***	-0.54***	-0.545***	-0.582***	-0.58***	-0.542***	-0.545***
	(0.034)	(0.04)	(0.04)	(0.042)	(0.033)	(0.039)	(0.039)	(0.041)
Eco/Geo-Demo/Ind/Gov	✓	✓	✓	✓	✓	✓	✓	✓
Municipal Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	11710	7745	7721	7302	11710	7745	7721	7302

Note: All continuous predictors are mean-centered and scaled by 1 standard deviation. *** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$.

4.6.2 Robustness Checks

The Tables 4.13 and 4.13 show the robustness estimates conducted on the votes for the ruling party, considering the average municipal share of families covered by the CI in the three-year period 2019-2021, on the voters and on the total resident population. The results obtained confirm the consistency of the general estimates presented and described in previous sections 4.4 and 4.5 since, in both tables, the impact of the average share of families receiving the monetary transfer on the electoral support for the incumbent party is positive and significant on the entire sample of Italian municipalities and considerably variable at a territorial level, between North and South. In the North, the impact of the income support programme remains negative and statistically significant, despite having a much stronger magnitude, in absolute values, than the analysis of the total of Italian municipalities. In South the impact of politics is positive and statistically significant, with the highest magnitudes.

In these robustness analyses, our research questions seem to be confirmed, with respect to the positive connection between obtaining the benefit and support for the incumbent (**Research Question $\mathbf{RQ}_1$**), especially by observing the Italian macro-regional spatial heterogeneity, which appears to be a determining driver to explain these connections (**Research Question $\mathbf{RQ}_2$**).

Also robustness analysis conducted considering the years of treatment individually, 2019, 2020, 2021, on the entire sample of Italian municipalities, and on the municipalities of Northern and Southern Italy (see Table 4.15) highlights results similar to those obtained in the general estimates, i.e. a positive and significant impact of the income support programme on votes in all Italian municipalities, negative and significant in the North and positive and significant in the South, Tables 4.4, 4.5, 4.6. These robustness estimates show interesting results from examining the magnitude of the coefficients of the independent variable specified in Eq. 39 in the different years of treatment. In fact, we observe a tendency towards a decrease in this magnitude, which indicates a weakening of the correlation - both negative and positive - between obtaining the benefit of the CI and the electoral support for the proposing party, during the period of observation of the policy. This data acquires particular importance in the context of the growth of multidimensional phenomena determined by the Covid-19 pandemic, a period during which the number of CI beneficiaries grew significantly. The described trend suggests that the adoption of the Italian Citizenship Income did not translate into a substantial electoral advantage for the IP in the long term.

Other robustness estimates are present in the Appendix . In particular, Table 4.1, which presents robustness analyses on Not IP considering the electoral series from 2013 to 2022, reports coefficients of the treatment variable consistent with those obtained in the other regressions that also considered 2008. The absence of a positive impact, in Italian municipalities, between CI and votes for Not IP confirms the hypothesis that sees the CI positively correlated to the consensus of the governing party in a different way (**Research question RQ**$_1$). Furthermore, also in these analyses, the difference in signs obtained between municipalities in the North and the South confirms the second fundamental question of this investigation (**Research question RQ**$_2$).

Heterogeneity in the direction of the coefficients is present in Tables 4.2 and 4.3, which report the robustness estimates comparing the pre-treatment years individually with the voting year in which the CI was present (2022). In particular, considering all Italian municipalities, the comparison between the votes obtained by the IP in 2018 and 2022 reports negative and significant coefficients of the independent variable, due to the high loss of consensus that the ruling party recorded between these two elections. Similarly, in the Northern municipalities, between 2008 and 2022, the sign of the coefficients for Not IP, both in the Chamber of Deputies and in the Senate, appear negative and significant. This evidence shows that even in these richer and more economically prosperous territories, which present fewer requests and needs for income support, the consensus of the Not IP parties has decreased.

Table 4.13: Robustness Checks. Votes for IP out of the Total Population (Italy, North and South)

	Italy		**North**		**South**	
	Chamber of Deputies	**Senate**	**Chamber of Deputies**	**Senate**	**Chamber of Deputies**	**Senate**
	Rob 1	**Rob 2**	**Rob 1**	**Rob 2**	**Rob 1**	**Rob 2**
Average share of households receiving CI of total families 2019-2021	0.587***	0.626***	-0.787***	-0.664***	0.85***	0.854***
	(0.029)	(0.028)	(0.068)	(0.067)	(0.052)	(0.051)
Eco/Geo-Demo/Ind/Gov	✓	✓	✓	✓	✓	✓
Municipal Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Time Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	16884	16884	11419	11419	5465	5465

Note: All continuous predictors are mean-centered and scaled by 1 standard deviation. *** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$.

Table 4.14: Robustness Checks. Votes for IP out of the Total Votes (Italy, North and South)

	Italy		North		South	
	Chamber of Deputies	Senate	Chamber of Deputies	Senate	Chamber of Deputies	Senate
	Rob 1	**Rob 2**	**Rob 1**	**Rob 2**	**Rob 1**	**Rob 2**
Average share of households receiving CI of total families 2019-2021	0.199***	0.299***	-0.491***	-0.366***	0.433***	0.465***
	(0.017)	(0.015)	(0.046)	(0.041)	(0.029)	(0.027)
Eco/Geo-Demo/Ind/Gov	✓	✓	✓	✓	✓	✓
Municipal Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Time Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	16884	16884	11419	11419	5465	5465

Note: All continuous predictors are mean-centered and scaled by 1 standard deviation. *** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$.

Table 4.15: Robustness Checks. Shares of Households Receiving CI in 2019-2021 (Italy, North and South)

Chamber of Deputies									
	Italy			**North**			**South**		
	2019	**2020**	**2021**	**2019**	**2020**	**2021**	**2019**	**2020**	**2021**
	Rob 1	**Rob 1**	**Rob 1**	**Rob 1**	**Rob 1**	**Rob 1**	**Rob 1**	**Rob 1**	**Rob 1**
Share of households receiving CI of total families 2019	0.43***			-0.89***			0.831***		
	(0.033)			(0.08)			(0.061)		
Share of households receiving CI of total families 2020		0.318***			-0.684***			0.553***	
		(0.023)			(0.06)			(0.041)	
Share of households receiving CI of total families 2021			0.179***			-0.31***			0.274***
			(0.013)			(0.033)			(0.021)
Eco/Geo-Demo/Ind/Gov	✓	✓	✓	✓	✓	✓	✓	✓	✓
Municipal Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	3733	3733	3733
Time Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	16884	16884	16884	11419	11419	11419	5465	5465	5465
Senate									
	Italy			**North**			**South**		
	2019	**2020**	**2021**	**2019**	**2020**	**2021**	**2019**	**2020**	**2021**
	Rob 1	**Rob 1**	**Rob 1**	**Rob 1**	**Rob 1**	**Rob 1**	**Rob 1**	**Rob 1**	**Rob 1**
Share of households receiving CI of total families 2019	0.467***			-0.778***			0.811***		
	(0.032)			(0.078)			(0.059)		
Share of households receiving CI of total families 2020		0.346***			-0.585***			0.544***	
		(0.023)			(0.059)			(0.04)	
Share of households receiving CI of total families 2021			0.194***			-0.249***			0.271***
			(0.013)			(0.033)			(0.021)
Eco/Geo-Demo/Ind/Gov	✓	✓	✓	✓	✓	✓	✓	✓	✓
Municipal Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	3733	3733	3733
Time Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	16884	16884	16884	11419	11419	11419	5465	5465	5465

Note: All continuous predictors are mean-centered and scaled by 1 standard deviation. *** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$.

4.7 Conclusions

This work explores the relationship between government income support programmes, regional disparities and voters' behaviour, leveraging the case of the Italian Citizenship Income initiative.

The analysis is based on a unique dataset that merges administrative data on programme beneficiaries with electoral data at the municipal level. Exploiting the municipal heterogeneity in the share of beneficiary households, the study employs a generalised difference-in-difference approach with continuous treatment to estimate the effect on the incumbent party vote share.

The results reveal that the income support programme positively influences citizens' voting behaviour in favour of the ruling party. This effect is more pronounced in disadvantaged regions, where the combination of poor institutional quality and high unemployment undermines development prospects. In such contexts, income support emerges as a strategic mechanism to strengthen support for the ruling party. However, under contrasting circumstances, income support policy has the opposite impact on voting behaviour.

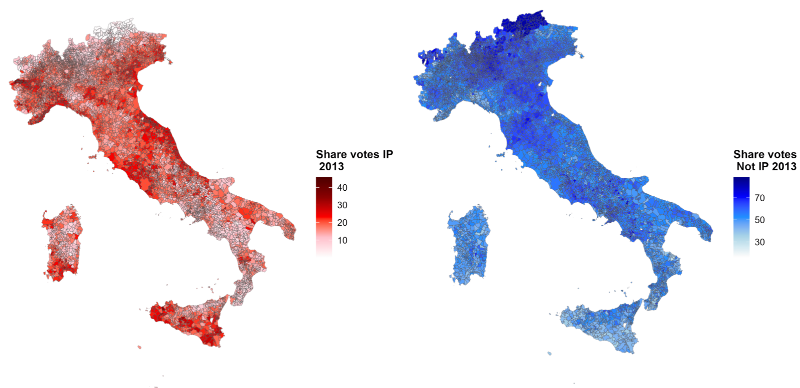
These findings add valuable insights to the larger discussions surrounding political economy, social choice, and public economics. They emphasise the significance of accounting for regional differences and contextual elements when evaluating the effectiveness of income support programmes in influencing political results. The research underscores the intricate relationship between income support policies and voter behaviour, accentuating the important role played by spatial and socioeconomic diversity in shaping the impact of policies on electoral outcomes. The findings not only advance academic understanding but also have practical implications for policymakers seeking to design effective anti-poverty initiatives that resonate with diverse regional contexts.

Appendix D

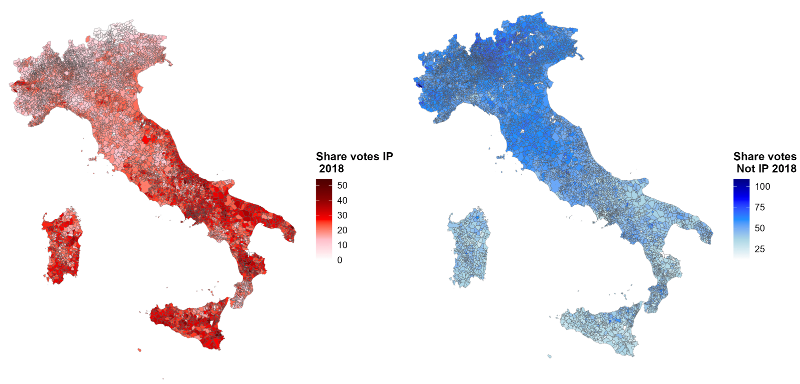
D1. Distribution of IP and Not IP Consensus

Figure D1: Distribution of Votes for IP and Not IP IN Chamber of Deputies

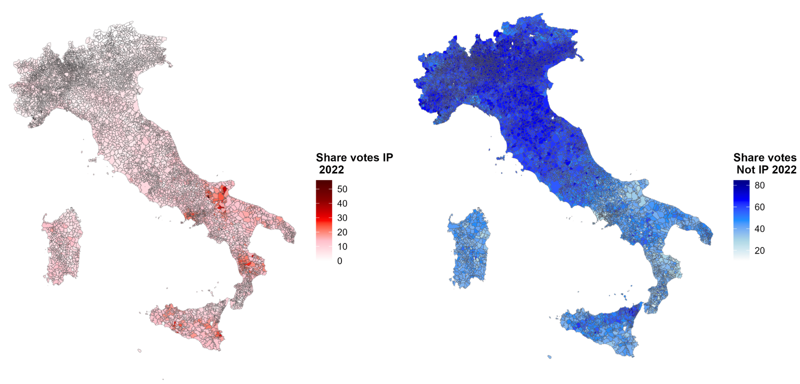
(a) 2013 - Left: IP, Right: Not IP



(b) 2018 - Left: IP, Right: Not IP



(c) 2022 - Left: IP, Right: Not IP



Source: Author's processing on Ministry of the Interior data.

D2. Other DiD Robustness Checks

Table 4.1: Robustness Checks. Votes for NotIP 2013-2022 (Italy, North and South)

	Italy		**North**		**South**	
	Chamber of Deputies	**Senate**	**Chamber of Deputies**	**Senate**	**Chamber of Deputies**	**Senate**
	Rob 1	**Rob 2**	**Rob 1**	**Rob 2**	**Rob 1**	**Rob 2**
Average share of households receiving CI of total families 2019-2021	-0.6***	-0.603***	0.593***	0.581***	-0.46***	-0.457***
	(0.025)	(0.025)	(0.074)	(0.074)	(0.046)	(0.045)
Eco/Geo-Demo/Ind/Gov	✓	✓	✓	✓	✓	✓
Municipal Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Time Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	16884	16884	11419	11419	5465	5465

Note: All continuous predictors are mean-centered and scaled by 1 standard deviation. *** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$.

Table 4.2: Robustness Checks. Results of the Two-Year Regression on the IP

	2013-2022					
	Italy		**North**		**South**	
	Chamber of Deputies	**Senate**	**Chamber of Deputies**	**Senate**	**Chamber of Deputies**	**Senate**
	Rob 1	**Rob 2**	**Rob 1**	**Rob 2**	**Rob 1**	**Rob 2**
Average share of households receiving CI of total families 2019-2021	0.974***	0.985***	-0.769***	-0.675***	0.833***	0.827***
	(0.025)	(0.024)	(0.083)	(0.081)	(0.044)	(0.042)
Eco/Geo-Demo/Ind/Gov	✓	✓	✓	✓	✓	✓
Municipal Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Time Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	11206	11206	7564	7564	3642	3642
	2018-2022					
	Italy		**North**		**South**	
	Chamber of Deputies	**Senate**	**Chamber of Deputies**	**Senate**	**Chamber of Deputies**	**Senate**
	Rob 1	**Rob 2**	**Rob 1**	**Rob 2**	**Rob 1**	**Rob 2**
Average share of households receiving CI of total families 2019-2021	-0.184***	-0.158***	-0.511***	-0.421***	0.297***	0.283***
	(0.019)	(0.019)	(0.06)	(0.059)	(0.036)	(0.036)
Eco/Geo-Demo/Ind/Gov	✓	✓	✓	✓	✓	✓
Municipal Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Time Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	11556	11556	7831	7831	3725	3725

Note: All continuous predictors are mean-centered and scaled by 1 standard deviation. *** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$.

Table 4.3: Robustness checks. Results of the Two-Year Regression on Not IP

2008-2022

	Italy		North		South	
	Chamber of Deputies	Senate	Chamber of Deputies	Senate	Chamber of Deputies	Senate
	Rob 1	Rob 2	Rob 1	Rob 2	Rob 1	Rob 2
Average share of households receiving CI of total families 2019-2021	-1.471***	-1.449***	-0.656***	-0.631***	-0.775***	-0.762***
	(0.027)	(0.027)	(0.074)	(0.076)	(0.061)	(0.061)
Eco/Geo-Demo/Ind/Gov	✓	✓	✓	✓	✓	✓
Municipal Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Time Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	11343	11343	7604	7604	3739	3739

2013-2022

	Italy		North		South	
	Chamber of Deputies	Senate	Chamber of Deputies	Senate	Chamber of Deputies	Senate
	Rob 1	Rob 2	Rob 1	Rob 2	Rob 1	Rob 2
Average share of households receiving CI of total families 2019-2021	-1.152***	-1.15***	0.954***	0.903***	-0.627***	-0.63***
	(0.035)	(0.035)	(0.11)	(0.11)	(0.068)	(0.065)
Eco/Geo-Demo/Ind/Gov	✓	✓	✓	✓	✓	✓
Municipal Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Time Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	11206	11206	7564	7564	3642	3642

2018-2022

	Italy		North		South	
	Chamber of Deputies	Senate	Chamber of Deputies	Senate	Chamber of Deputies	Senate
	Rob 1	Rob 2	Rob 1	Rob 2	Rob 1	Rob 2
Average share of households receiving CI of total families 2019-2021	-0.182***	-0.192***	0.344***	0.357***	-0.349***	-0.337***
	(0.023)	(0.022)	(0.068)	(0.067)	(0.051)	(0.05)
Eco/Geo-Demo/Ind/Gov	✓	✓	✓	✓	✓	✓
Municipal Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Time Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	11556	11556	7831	7831	3725	3725

Note: All continuous predictors are mean-centered and scaled by 1 standard deviation. *** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$.

Bibliography

Accetturo, A., Albanese, G., Torrini, R., Depalo, D., Giacomelli, S., Messina, G., Scoccianti, F., and Vacca, V. P. (2022). Il divario nord sud: sviluppo economico e intervento pubblico. In *Bank of Italy Workshops and Conferences Series (2022)*.

Acciari, P. and Mocetti, S. (2013). The geography of income inequality in italy. Technical report, Bank of Italy, Economic Research and International Relations Area.

Alabrese, E., Becker, S. O., Fetzer, T., and Novy, D. (2019). Who voted for brexit? individual and regional data combined. *European Journal of Political Economy*, 56:132–150.

Albanese, G., Barone, G., and de Blasio, G. (2022). Populist voting and losers' discontent: Does redistribution matter? *European Economic Review*, 141:104000.

Albanese, G., de Blasio, G., and Incoronato, L. (2024). Government transfers and votes for state intervention. *American economic journal: economic policy*, 15:20.

Alesina, A. and Fuchs-Schündeln, N. (2007). Good-bye lenin (or not?): The effect of communism on people's preferences. *American Economic Review*, 97(4):1507–1528.

Alesina, A. and Perotti, R. (1996). Income distribution, political instability, and investment. *European economic review*, 40(6):1203–1228.

Alfano, V. and Ercolano, S. (2021). Social capital, quality of institutions and lockdown. evidence from italian provinces. *Structural Change and Economic Dynamics*, 59:31–41.

Angelucci, D., Trastulli, F., and Flumeri, E. (2022). Tiene la correlazione reddito di cittadinanza-voto al m5s.

Araújo, V. (2021). Do anti-poverty policies sway voters? evidence from a meta-analysis of conditional cash transfers. *Research & Politics*, 8(1).

Arbia, G. (2023). Income inequality, redistribution and economic growth. *SpringerBriefs in Economics*.

Atkinson, A. B. (2015). *Inequality: What can be done?* Harvard University Press.

Atkinson, A. B. and Piketty, T. (2007). *Top incomes over the twentieth century: a contrast between continental european and english-speaking countries*. oup Oxford.

Baldini, M. and Gori, C. (2019). Il reddito di cittadinanza. *il Mulino*, 68(2):269–277.

Basile, L. (2023). Political equality in unequal territories: the impact of territorial inequalities on voter turnout in italy's regional elections. *Regional & Federal Studies*, pages 1–25.

Blanchard, O. and Rodrik, D. (2023). *Combating Inequality: Rethinking Government's Role*. MIT press.

Bloise, F., Chironi, D., della Porta, D., and Pianta, M. (2023). Inequality and elections in italy, 1994–2018. *Italian Economic Journal*, pages 1–23.

Bloise, F., Chironi, D., and Pianta, M. (2021). Inequality and voting in italy's regions. *Territory, Politics, Governance*, 9(3):365–390.

Busilacchi, G. and Fabbri, A. (2023). Does the unemployment trap still exist? the case of the italian minimum income scheme. *Social Policy & Administration*.

Callaway, B., Goodman-Bacon, A., and Sant'Anna, P. H. (2024). Difference-in-differences with a continuous treatment. Technical report, National Bureau of Economic Research.

Callaway, B. and Huang, W. (2020). Distributional effects of a continuous treatment with an application on intergenerational mobility. *Oxford Bulletin of Economics and Statistics*, 82(4):808–842.

Card, D. (1992). Using regional variation in wages to measure the effects of the federal minimum wage. *Ilr Review*, 46(1):22–37.

Carreras, M. (2019). 'what do we have to lose?': Local economic decline, prospect theory, and support for brexit. *Electoral Studies*, 62:102094.

Carta, F. and De Philippis, M. (2021). The impact of the covid-19 shock on labour income inequality: Evidence from italy. *Bank of Italy Occasional Paper*, (606).

Caselli, M., Fracasso, A., and Traverso, S. (2020). Globalization and electoral outcomes: Evidence from italy. *Economics & Politics*, 32(1):68–103.

Cerda, R. and Vergara, R. (2008). Government subsidies and presidential election outcomes: evidence for a developing country. *World Development*, 36(11):2470–2488.

Cerqua, A., Ferrante, C., and Letta, M. (2023). Electoral earthquake: Local shocks and authoritarian voting. *European Economic Review*, 156:104464.

Cerulli, G. (2015). ctreatreg: Command for fitting dose–response models under exogenous and endogenous treatment. *The Stata Journal*, 15(4):1019–1045.

Cerulli, G. and Ventura, M. (2021). A dose–response approach to evaluate the effects of different levels of partial credit guarantees. *Applied Economics*, 53(12):1418–1434.

Checchi, D., Dachille, G. P., and De Paola, M. (2021). Reddito di cittadinanza, caratteristiche socio-economiche e capitale sociale (citizens' income, socio-economic characteristics, and social capital). *Politica economica*, 37(1):121–148.

Chelli, F. M., Ermini, B., Gallegati, M., and Gentili, A. (2023). Investigating regional disparities in italy's well-being since unification (1871–2011). *Italian Economic Journal*, 9(2):697–722.

Chiaramonte, A., Emanuele, V., Maggini, N., and Paparo, A. (2022). Radical-right surge in a deinstitutionalised party system: The 2022 italian general election. *South European Society and Politics*, 27(3):329–357.

Coco, G. and De Vincenti, C. (2020). *Una questione nazionale: il Mezzogiorno da "problema" a "opportunità"*. Società editrice Il mulino.

Conover, E., Zárate, R. A., Camacho, A., and Baez, J. E. (2020). Cash and ballots: Conditional transfers, political participation, and voting behavior. *Economic Development and Cultural Change*, 68(2):541–566.

Corradini, C. (2021). Local institutional quality and economic growth: A panel-var analysis of italian nuts-3 regions. *Economics Letters*, 198:109659.

Cox, G. W. (2009). 13 swing voters, core voters, and distributive politics. *Political representation*, 342.

De Chaisemartin, C. and d'Haultfoeuille, X. (2020). Two-way fixed effects estimators with heterogeneous treatment effects. *American economic review*, 110(9):2964–2996.

De La O, A. L. (2013). Do conditional cash transfers affect electoral behavior? evidence from a randomized experiment in mexico. *American Journal of Political Science*, 57(1):1–14.

De Luca, G., Lisi, D., Martorana, M., and Siciliani, L. (2021). Does higher institutional quality improve the appropriateness of healthcare provision? *Journal of Public Economics*, 194:104356.

Dijkstra, L., Poelman, H., and Rodríguez-Pose, A. (2020). The geography of eu discontent. *Regional Studies*, 54(6):737–753.

D'Haultfœuille, X., Hoderlein, S., and Sasaki, Y. (2023). Nonparametric difference-in-differences in repeated cross-sections with continuous treatments. *Journal of Econometrics*, 234(2):664–690.

D'Ingiullo, D. and Evangelista, V. (2020). Institutional quality and innovation performance: evidence from italy. *Regional studies*, 54(12):1724–1736.

Elinder, M., Jordahl, H., and Poutvaara, P. (2015). Promises, policies and pocketbook voting. *European Economic Review*, 75:177–194.

Eurostat (2019). *Eurostat regional yearbook – 2019 edition*. Publications Office.

Faggian, A., Modica, M., Modrego, F., and Urso, G. (2021). One country, two populist parties: Voting patterns of the 2018 italian elections and their determinants. *Regional Science Policy & Practice*, 13(2):397–413.

Felice, E. (2018). Regional income inequality in italy in the long run (1871–2010): Patterns and determinants. In *The Economic Development of Europe's Regions*, pages 177–203. Routledge.

Ferrara, A. R. and Nisticò, R. (2019). Does institutional quality matter for multidimensional well-being inequalities? insights from italy. *Social Indicators Research*, 145(3):1063–1105.

Filippetti, A. and Cerulli, G. (2018). Are local public services better delivered in more autonomous regions? evidence from european regions using a dose-response approach. *Papers in Regional Science*, 97(3):801–827.

Fiorina, M. P. (1978). Economic retrospective voting in american national elections: A micro-analysis. *American Journal of political science*, pages 426–443.

Gallo, G. (2021). Regional support for the national government: Joint effects of minimum income schemes in italy. *Italian Economic Journal*, 7(1):149–185.

Gallo, G. and Raitano, M. (2023). Sos incomes: Simulated effects of covid-19 and emergency benefits on individual and household income distribution in italy. *Journal of European Social Policy*, 33(1):101–116.

Gethin, A., Martínez-Toledano, C., and Piketty, T. (2022). Brahmin left versus merchant right: Changing political cleavages in 21 western democracies, 1948–2020. *The Quarterly Journal of Economics*, 137(1):1–48.

Girardi, E. (2023). Precarious work and challenger parties: how precarity influenced vote choice in the 2018 italian election. *Italian Political Science Review/Rivista Italiana di Scienza Politica*, pages 1–19.

Giuliani, M. (2023). Did the citizenship income scheme do it? the supposed electoral consequence of a flagship policy. *Italian Political Science Review/Rivista Italiana di Scienza Politica*, pages 1–9.

Gromadzki, J., Sałach, K., and Brzeziński, M. (2024). When populists deliver on their promises: the electoral effects of a large cash transfer programme in poland. *Economica*, 91(361):320–345.

Hansen, C. B. (2007). Generalized least squares inference in panel and multilevel models with serial correlation and fixed effects. *Journal of econometrics*, 140(2):670–694.

Imai, K. and Kim, I. S. (2021). On the use of two-way fixed effects regression models for causal inference with panel data. *Political Analysis*, 29(3):405–415.

Imami, D., Pojani, D., and Merkaj, E. (2023). Electoral politics of disaster: how earthquake and pandemic relief was used to earn votes. *Cambridge Journal of Regions, Economy and Society*, 16(1):93–104.

INPS (2022a). Conoscere il paese per costruire il futuro. XXI Rapporrto Annuale. INPS, Roma.

INPS (2022b). XXI Rapporto Annuale: Relazione del Presidente. INPS, Roma.

INPS (2023). XXII Rapporto annuale. INPS, Roma.

Ippolito, M., Ercolano, S., and Cicatiello, L. (2023). You want me, but i no longer want you. the punishment-reward mechanism and institutional quality in italian regions. *Contemporary Italian Politics*, pages 1–16.

Lalioti, V. (2016). The curious case of the guaranteed minimum income (gmi): Highlighting greek 'exceptionalism'in a southern european context. *Journal of European Social Policy*, 26(1):80–93.

Levi, E., Mariani, R. D., and Patriarca, F. (2020). Hate at first sight? dynamic aspects of the electoral impact of migration: the case of ukip. *Journal of Population Economics*, 33:1–32.

Levi, E. and Patriarca, F. (2020). An exploratory study of populism: the municipality-level predictors of electoral outcomes in italy. *Economia politica*, 37:833–875.

Levitt, S. D. and Snyder Jr, J. M. (1997). The impact of federal spending on house election outcomes. *Journal of political Economy*, 105(1):30–53.

Lewis-Beck, M. S. (1985). Pocketbook voting in us national election studies: fact or artifact? *American Journal of Political Science*, pages 348–356.

Lewis-Beck, M. S. and Stegmaier, M. (2000). Economic determinants of electoral outcomes. *Annual review of political science*, 3(1):183–219.

Lewis-Beck, M. S. and Stegmaier, M. (2008). The economic vote in transitional democracies. *Journal of Elections, Public Opinion and Parties*, 18(3):303–323.

Lewis-Beck, M. S. and Stegmaier, M. (2018). Economic voting. *The Oxford handbook of public choice*, 1:247–265.

Loeffler, H. (2023). Does a universal basic income affect voter turnout? evidence from alaska. *Political Science Research and Methods*, 11(3):521–536.

Maitino, M. L., Mariani, M., Patacchini, V., Ravagli, L., and Sciclone, N. (2024). The employment effects of the italian minimum guaranteed income scheme reddito di cittadinanza. *Italian Economic Journal*, pages 1–33.

Manacorda, M., Miguel, E., and Vigorito, A. (2011). Government transfers and political support. *American Economic Journal: Applied Economics*, 3(3):1–28.

Marx, P. (2016). The insider-outsider divide and economic voting: Testing a new theory with german electoral data. *Socio-Economic Review*, 14(1):97–118.

Meyer, B. D. (1995). Natural and quasi-experiments in economics. *Journal of business & economic statistics*, 13(2):151–161.

Monturano, G. (2023). *Comunicare e percepire le politiche economiche e fiscali. L'asimmetria informativa e comunicativa nelle misure di contrasto alla povertà e nelle imposte sui patrimoni*, pages 257–269.

Monturano, G., Sonzogno, G. V., and Resce, G. (2023). La capacità di target territoriale del reddito di cittadinanza. *Menabò. Etica ed Economia*, pages 1–10.

Mosconi, F. and D'Ingiullo, D. (2023). Institutional quality and innovation: evidence from emilia-romagna. *Economics of Innovation and New Technology*, 32(2):165–197.

Nifo, A. and Vecchione, G. (2014). Do institutions play a role in skilled migration? the case of italy. *Regional Studies*, 48(10):1628–1649.

Palomino, J. C., Rodríguez, J. G., and Sebastian, R. (2020). Wage inequality and poverty effects of lockdown and social distancing in europe. *European economic review*, 129:103564.

Pierson, P. (1996). The new politics of the welfare state. *World politics*, 48(2):143–179.

Pignataro, G. and Prarolo, G. (2020). Learning, proximity and voting: theory and empirical evidence from nuclear referenda. *Social Choice and Welfare*, 55:117–147.

Piketty, T. (2018). Brahmin left vs merchant right: Rising inequality & the changing structure of political conflict.

Piketty, T. and Saez, E. (2014). Inequality in the long run. *Science*, 344(6186):838–843.

Pop-Eleches, C., Pop-Eleches, G., et al. (2012). Targeted government spending and political preferences. *Quarterly Journal of Political Science*, 7(3):285–320.

Raitano, M., Jessoula, M., Gallo, G., Pagnini, C., et al. (2021). *Fighting poverty and social exclusion-including through minimum income schemes*. European Parliament.

Rodríguez-Pose, A. (2018). The revenge of the places that don't matter (and what to do about it). *Cambridge journal of regions, economy and society*, 11(1):189–209.

Rodríguez-Pose, A., Lee, N., and Lipp, C. (2021). Golfing with trump. social capital, decline, inequality, and the rise of populism in the us. *Cambridge Journal of Regions, Economy and Society*, 14(3):457–481.

Rovny, A. E. and Rovny, J. (2017). Outsiders at the ballot box: operationalizations and political consequences of the insider–outsider dualism. *Socio-Economic Review*, 15(1):161–185.

Tonutti, G., Bertarelli, G., Giusti, C., and Pratesi, M. (2022). Disaggregation of poverty indicators by small area methods for assessing the targeting of the "reddito di cittadinanza" national policy in italy. *Socio-Economic Planning Sciences*, 82:101327.

Zucco Jr, C. (2013). When payouts pay off: Conditional cash transfers and voting behavior in brazil 2002–10. *American journal of political science*, 57(4):810–822.