Classe di Scienze

Corso di perfezionamento in
Data Science

34 ciclo

Anno accademico 2022/23

Between Two Crises: Local-level Analysis of Poverty and Social Exclusion in Italy

SECS-S/05

Candidato
dr. Giovanni Tonutti

Relatori
prof.ssa Monica Pratesi
dr.ssa Gaia Bertarelli
prof. Mauro-Sylos Labini

Supervisore
prof.ssa Monica Pratesi

To Elio

"Unwinding ourselves from our neighbors' deprivation and refusing to live as enemies of the poor will require us to pay a price. It's the price of our restored humanity and renewed country."

Matthew Desmond

Acknowledgments

The completion of this doctoral thesis represents undoubtedly the most challenging intellectual endeavour I set out to achieve so far. This work would not have been possible without the mentorship of my supervisor, prof. Monica Pratesi, whose support extended far beyond the supervision of my research activities, providing me with the opportunity to grow as a researcher and analyst. My deepest and most sincere gratitude goes to Dr. Gaia Bertarelli. Without her pragmatism, determination, and genuine interest in my work, the completion of this research would not have been possible, especially since taking up my current full-time job. Thanks should also go to my co-author prof. Mauro Sylos Labini for his patience in leading me with uncompromising rigour through the fascinating discipline of causal inference methods. I am grateful to prof. Roventini for facilitating access to data on RdC beneficiaries from INPS, without which this research would not have happened.

My years in Pisa, marked by the unnerving uncertainty cast upon us by the pandemic, would not have been the same without the community and comradeship of the eXploit collective. In a time of deep isolation, the assemblies, rallies, and demonstrations not only provided the ideological compass for guiding my research efforts but allowed me to feel part of a community at a time when most other human interactions were lost.

The pursuit of this doctoral degree coincided with a period of deep personal grief and transformation, which I shared, despite the distance, with my brother and my mother. I am endlessly grateful for her strenuous support and care towards my endeavour, demonstrated countless of times, right from the day of my selection interview. I want to thank Paaras for understanding my time in London had come to an end and for supporting my move to Pisa. Finally, a thank you to Lotti. We can be rocks or we can be leaves, together or apart.

Contents

List of Figures

List of Tables

Summary

This dissertation reviews the evolving landscape regarding poverty and the tools to address it in Italy during the period of deep economic crisis following the global Great Recession of 2007/2008. The three core chapters of this dissertation investigate how pre-existing geographical inequalities shape the evolution of poverty-related aspects, affect the policy responses enacted to contrast them and impact the political consequences of such policies. The novelty of this work lies in providing an account of these recent developments at an unprecedented level of geographical detail. Each chapter looks at a different regional and local aspect related to poverty and social exclusion, through the application of statistical and econometrics models to a mix of survey and administrative data. Chapter 1 introduces the topic by briefly reviewing the evolution of poverty studies in Italy and putting recent macroeconomic dynamics into context. It highlights the relevance of local-level analyses in a global setting of rising regional inequalities as well as discusses some of the challenges this type of approach faces in relation to the sources of data available. Through the lenses of economic geography and industrial change, Chapter 2 analyzes how the phenomenon of in-work poverty has evolved in Italy in the period between 2008 and 2017 and explores the links between the rise of low-pay work and processes of de-industrialisation and local shifts in sectoral employment. From there, we move towards an analysis of the political economy with a focus on policy and politics. Chapter 3 provides, to the best of our knowledge, the first targeting assessment at the small-area level of the most extensive anti-poverty program ever introduced in the country, the Citizen Income (RdC). Chapter 4 looks at the political consequences of the introduction of the RdC by assessing its impact on the electoral outcomes of the Five Star Movement (M5S), the political party which introduced the program.

The chapters provide the following key insights. First, while the incidence of low pay and in-work poverty increased significantly across the whole country, a local area analysis highlights how a decade characterised by deep economic recession and slow recovery has contributed to widening the country's north-south dualism in relation to the phenomenon

of low pay. Second, we show that this process is linked to specific shifts in the local industrial composition and in particular to the process of de-industrialisation. Third, such regional and local disparities are reflected in the strong geographical heterogeneity of the targeting performance of the RdC, the anti-poverty policy rolled out across the country in 2019. Fourth, in a context of increasing political polarization on the issue of poverty, features (and flaws) inherent to the design and implementation of the RdC policy directly affected the political consequences of the measure. Initially, the measure was met with widespread disappointment by its beneficiaries, who gradually rallied behind the incumbent under the threat that the main opposition party might suspend the program.

These findings have implications for both academic literature and policy-making. We provide novel evidence of the geographical distribution of in-work poverty across non-administrative small area boundaries and explore its link with sectoral employment dynamics amidst a context of de-industrialisation. Moreover, in providing the very first assessment of the targeting of the most ambitious social welfare program yet to be introduced in Italy, the analysis provides tangible suggestions on how the policy could be improved to ensure it reaches a wider cohort of households living in poverty. Lastly, this work represents a one-of-a-kind case study to understand the political consequences of government cash transfers in the context of a highly partisan measure implemented by a populist government. Together, these findings contribute robust quantitative evidence to the expanding body of literature concerned with studying how poverty became a persistent scourge in Italian society.

Keywords: Poverty, social statistics, small area analysis, in-work poverty, targeting, political economy.

Chapter 1

Introduction

The very first attempt at measuring the phenomenon of poverty dates back to 1893 when the British merchant Charles Booth published the 17-volumes *Life and Labour of the people in London*. More broadly, the field of poverty studies emerged in the late 19th century in response to the vast and increasingly visible disparities of industrial capitalism in Western Europe and the United States (O'Connor, 2016). In Italy, however, the issue has long remained neglected from both the public discourse and academic focus. With the exception of a short-lived parliamentary commission established in 1951, the phenomenon of poverty was essentially ignored for the following 30 years(Saraceno and Benassi, 2020). The field of poverty studies gradually developed in the late 1980s, driven by two factors. On the one hand, slower economic growth, which later became prolonged stagnation, made evident the structural social problems left unresolved by the post-World War II economic development. On the other hand, the remarkable progress in theoretical elaboration and the availability of new sources of microeconomic data have greatly stimulated research in the social sciences (Brandolini, 2021).

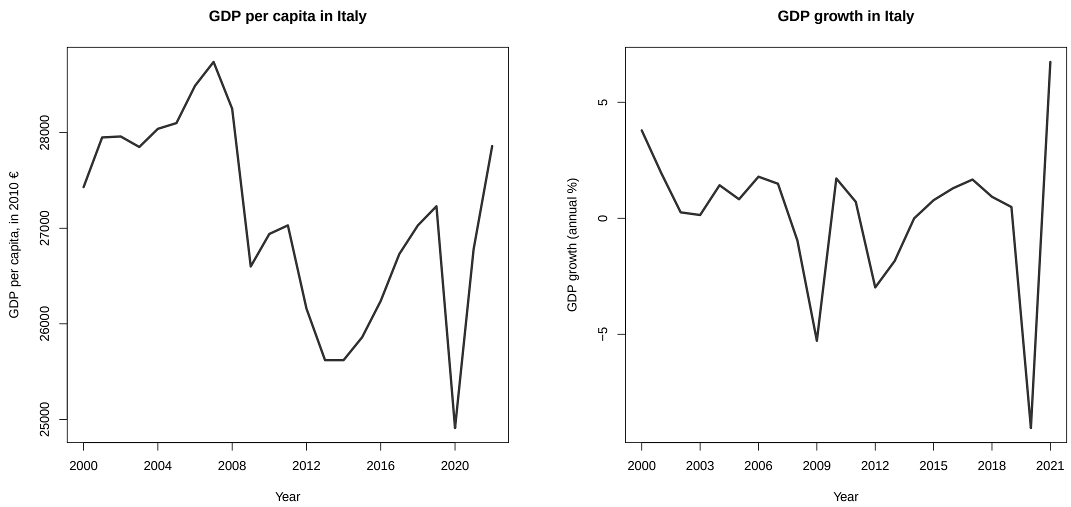
It is within this context that the research work presented herewith is situated. Indeed, the trend of slow economic growth which emerged in Italy in the early 1990s consolidated further at the turn of the new millennium, where Italy took the role of the "new sick man of Europe" (The Economist, 2005). In the years that followed, the country's economy was one of the hardest hit by the 2007/2008 financial crisis and subsequent sovereign debt crisis, and one of the slowest in recovering from it before entering the Covid-19 pandemic phase. During the same period, national indicators pointing to poverty have steadily increased with no signs of a reversal in trend at the point of writing. In this context, the issue not only became the subject of interest of an increasing number of research publications, but it also rose to the forefront of public debate. In 2017 and 2019, for the first time in the country's history, two successive governments introduced large national programs aimed at tackling and reducing poverty. In the context of political instability, characterising Italian politics, such interventions soon became the subject of political polarization to the extent that poverty and the tools to contrast it featured as some of the most salient issues in the country's two latest general elections. At the same time, (1) the increasing availability of rich and detailed data sources, (2) the development of increasingly robust statistical and econometrics methods and (3) mature conceptualisations of poverty and its nuances allow for the investigation of these phenomena at unprecedented levels of detail. Among such conceptualisations, Saraceno et al. (2022) provide a useful definition of a poverty regime as the specific combination of labour market conditions, the balance between public and private (family) responsibility in offering social protection and social and cultural norms around the labour division in households. This work does not have the ambition of providing an exhaustive compendium of the phenomenon of poverty in Italy and no overarching definition of poverty is adopted in this work. Instead, various indicators related to socio-economic deprivation are presented and discussed separately in each chapter. Building on the vast literature on labour economics, regional industrial policy, poverty targeting and political economy, the research presented in the following chapters touches upon most of the pillars of the poverty regime that characterised the Italian context in the years between 2008 and 2019. In doing so, it focuses on how dynamics related to poverty and social exclusion reflect some of the country's characteristics in relation to its socio-economic geographical tissue.

The remainder of Chapter 1 will provide a background of the main socio-economic trends that characterised the period between 2008 and 2019, while discussing the relevance of local-level analysis in the context of increasing regional inequalities and presenting the state-of-the-art with regard to data availability in the context of poverty studies. Chapter 2 presents a small area-level analysis of the distribution of in-work poverty at the level of Local Labour Systems. It then combines administrative data on local industrial composition to explore how the phenomenon of low-pay is linked with trends of de-industrialisation. Chapter 3 focuses on the recent introduction of Italy's largest anti-poverty program, the Citizen's Income or in Italian Reddito di Cittadinanza (RdC). By applying small area estimation models to map the distribution of relative and absolute poverty at the level of the three degrees of urbanisation in each of Italy's 20 regions, it provides an assessment of the RdC targeting in relation to the cohort of households in poverty. Chapter 4 focuses on the political acceptability of the same measure. Building on a panel data set of municipality-level election data, it studies the extent to which the introduction of the RdC impacted the electoral outcomes of the Five Star Movement (M5S), the party that designed and introduced this exact policy.

1.1 Setting the scene - Italy's economic decline

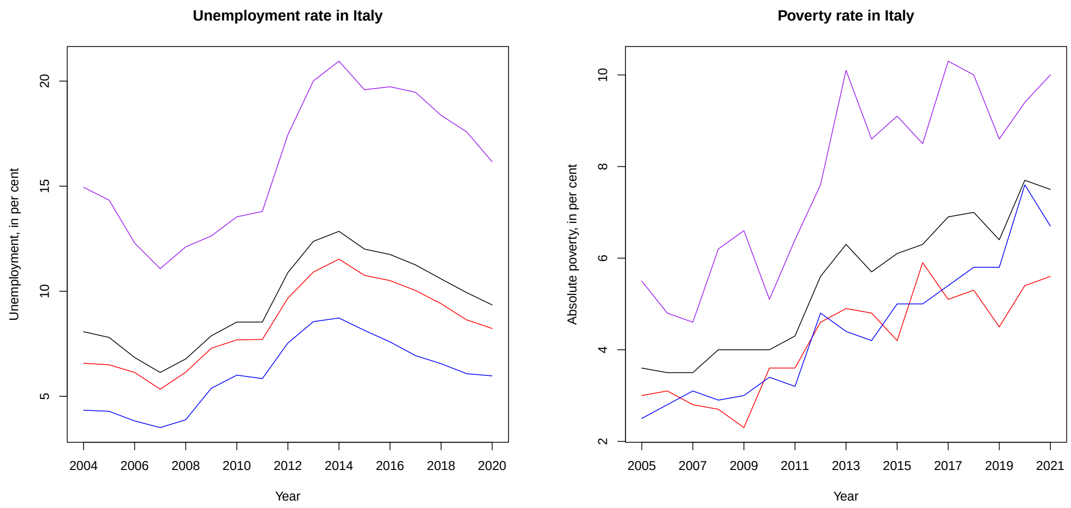
Italy's economy has no doubt been one of the hardest hit by the financial and sovereign debt crisis of the years 2008 and 2011 and to date, many of its macroeconomic indicators still have to recover to the pre-crisis levels (DIPE, 2022). The reasons for such a severe downturn and slow recovery are likely to be found in the period of sluggish economic growth the country experienced in the two decades predating the Great Recession. Since the 1990s, the country entered a phase of creeping crises of stagnation and inflation, with a progressive loss of ground in terms of competitiveness and per-capita income compared to other European economies (Manasse, 2013). At the turn of the new millennium, economists and commentators referred to the country as the new "sick man of Europe" (The Economist, 2005). Low productivity has been identified as the key driver for Italy's stagnating economic performance (Daveri et al., 2005) with total factor productivity (TFP) growth already stalling starting as early as the 1970s. For years, however, the country's economy, heavily relying on manufacturing and particularly on labour-intense sectors such as textiles, furniture, machine tools, food-processing and white goods, leveraged monetary policy and currency devaluation as tools to maintain economic competitiveness. With the turn to the 1990s, two main drivers seem to have accelerated the process of economic stagnation the country had undertaken (D'ippoliti and Roncaglia, 2011). First, Italy's adoption of the Euro meant monetary policy and currency devaluation was no longer a policy option in the quest to maintain economic competitiveness. This was reflected in a generalised loss of competitiveness in many industrial sectors further aggravated by increasing competition from Eastern Europe and China. Second, with the aim of reducing historically higher than-regional-average unemployment rates, the country pursued a series of reforms aimed at liberalising the labour market. While the share of employment increased consistently throughout the 1990s and the first years of the 2000s, especially among the female workforce, econometric evidence suggests how the introduction of more flexibility in the labour market, especially in relation to the ability of firms to hire workers on atypical work contracts, is linked to lower rates of labour productivity growth (Lucidi and Kleinknecht, 2010). As a result of this period of modest employment and productivity growth, Italy had in fact already entered a recession before the global crisis of 2008. In seasonally adjusted terms, the decline in investment began in the last quarter of 2006 and that of exports in the first quarter of 2007, soon followed by imports. Consumption (public and private) remained relatively constant throughout the period. As a result, in the fourth quarter of 2007, Italy suffered a drop in seasonally adjusted real GDP by 0.4% on a quarterly basis (corresponding to -1.6% on an annual basis) while the Eurozone recorded the first drop in production only in the second trimester of 2008 (EUROSTAT, 2022b), as shown in Figure 1.1, depicting GDP per capita and GDP growth over time in Italy.

Figure 1.1: GDP per capita at 2010 constant prices (left panel) and GDP growth (right). Source: own elaboration based on EUROSTAT and World Bank data.



While Italian banks were less exposed to the global financial contagion of the 2008 subprime mortgage crisis, the country nevertheless entered a period of recession as the result of the overall global downturn. The Great Recession of 2008 was quickly followed by the sovereign debt crisis which put Italy right in the eye of the storm. To quell investors' fears and financial turmoil, the country committed to a roadmap of tight fiscal policy measures aimed at reducing the deficit to GDP ratio and overall public debt. As a result, public and private investment fell from 3.2 to 2.3% and from 10 to 8% of GDP, respectively. In 2019, the index of industrial production was 20 points lower than the 2000 benchmark year and showed only slight trends of recovery from the 2008 slump (DIPE, 2022). The effect of such a severe downturn and subsequent sluggish economic recovery is reflected in multiple other socio-economic indicators. During the same period of observation, the share of Italian households living in absolute poverty as measured by ISTAT reached the 6.4% of the total population in the year 2019 from 4% in 2008 (ISTAT, 2022a). Similarly, unemployment close to doubled from 6.7% in 2008 to 10.5% in 2017 (EUROSTAT, 2021), while the share of workers hired through temporary employment contracts rose from 13.3 to 15.4%. Both the unemployment and the absolute poverty rate are shown below in Figure 1.2, showing national averages over time as well as averages for the South, North and Center of Italy. Finally, average real wages in Italy have actually decreased by approximately more than 1% since 2007, after accounting for inflation (OECD.Stat, 2022).

Figure 1.2: Unemployment rate (left panel) and absolute poverty rate (right panel). Black refers to the national average, purple refers to South, blue to North and red to Center. Source: own elaboration based on ISTAT.



It is important to notice how this first period of economic stagnation, and the subsequent downturn did not impact the country homogeneously. In a context of strong economic dualism tracing back to the very formation of the modern Italian state, with Northern and Southern areas on different paths of industrial development, the economic setbacks experienced in the last two decades affected the country's macro areas differently. Given a historically undersized private sector, the public sector always carried more economic weight and relevance in the South than in the rest of the country. The Southern economy has thus found itself particularly exposed over the past decade to the austerity measures imposed by the sovereign debt crisis, which has resulted in declining employment in public administration and reduced public investment, on which many private sector activities also indirectly depended (Torrini et al., 2022). As a result, the ratio between the per capita GDP of the South with that of the Center-North dropped from 58% in 2008 to 55% in 2019 (De Philippis et al., 2022), while the rate of unemployment in the South raised from 11 to 15% in the period 2008-2019, with a peak of 21% in 2014. The impact of these macro-economics trends has been reflected in the living standards of households. The gap in the share of households living in absolute poverty between the country's macro areas went from only 1.5 p.p. in 2007 (4.5% in the South via-à-vis 3% in the North and Centre) to 4.5 p.p. in 2018, with an overall incidence in the South almost twice as large as in the other areas of the country (ISTAT, 2022a).

1.2 Local level analysis and geographical inequalities

Amidst a backdrop of increasing geographical inequality and the presence of strong economic dualism, the regional and local analyses assume important relevance. This not only applies to the Italian context but is extendable, as a minimum, to most industrialised societies in which processes of de-industrialisation set forth a trend of widening regional divides. The diffusive processes of economic development which characterised the post-World War II period have in fact weakened in the last three to four decades and the gap between peripheral regions and areas boasting urban centres capable of developing strong agglomeration economies widened (Barca et al., 2012; Gräbner and Hafele, 2020). This is a process driven by structural economic changes and by the development of advanced service activities, most frequently located in urban areas, which have replaced industrial production, affected by pervasive automation processes or relocated to countries characterized by lower labour costs, with lower knowledge content (Benito and Ezcurra, 2005). In this context of rising regional inequality, to capture and describe this trend at an ever-more granular level of geographical details is not only necessary to establish a nuanced and accurate description of the socio-economic reality, but it provides the evidence basis for targeted public policy interventions. Without pragmatic 'place-sensitive development approaches' (Iammarino et al., 2019), indeed, dynamics of geographical inequalities have proven to lead to increasing divergence and so-called 'spatial traps'. In a study of 283 regions across 32 European countries, Pinheiro et al. (2022) unveil how the process of economic diversification and complexity spurred by innovation leads only the most economically advanced regions to diversify into highly complex activities while lagging regions focus on low value-added activities, creating a spatial inequality feedback loop. The implications of these dynamics extend well beyond economic indicators to encompass all aspects of society. In particular, the recent rise of populist political forces and leaders across Western democracies, including Italy, and events such as the Brexit referendum shifted public attention over the 'geography of discontent' (see, e.g., Mc-

Cann, 2020; Dijkstra et al., 2020; Rodríguez-Pose, 2018). This paradigm highlighted how voters from 'left behind' areas expressed their discontent through the support of candidates explicitly running on anti-establishment platforms (see, for instance, Colantone et al., 2021, on the impact of trade exposure). Such discontent reflects voters' experiences of relative economic decline, reinforced by stagnating real wages and austerity policies introduced since the global financial crisis of 2008 (MacKinnon et al., 2022). In this context, social science research focusing on the geospatial dynamics of socio-economic phenomena assumes further relevance. A distributional approach at the regional and local level not only provides evidence for how interventions should be targeted but can also lay the ground for assessing the public acceptability of such interventions, a crucial aspect in the attempt of reducing geographical inequalities. In doing so, it is important to consider how the dynamics described so far often unravel well beyond traditional administrative boundaries, making the choice of the unit of analysis a crucial aspect (Isserman et al., 1987). Administrative boundaries, indeed, are typically established via a top-down approach, with the primary aim to establish the perimeters within which sub-national institutions exercise their remit. Though in some cases this process typically takes into account the structure of the social and economic realities, it is slow to adapt to how such realities evolve (Casado-Díaz et al., 2011). Relying on administrative boundaries to capture the geographical distribution of socio-economic phenomena, therefore, can result in misrepresentation, especially when such boundaries contain highly heterogeneous realities (ISTAT, 2014). The availability of detailed geo-spatial information has, in recent years, given rise to the production of non-administrative boundaries and typologies or so-called functional regions (Karlsson and Olsson, 2006). Contrary to the notion of administrative geographies, these units result from the spontaneous organisation of the actions of socio-economic subjects and thus better approximate the perimeters of networks, exchanges and flows that characterize place (Farmer and Fotheringham, 2011). In the context of this research, for example, metrics of poverty and social exclusions will be produced at geographical units capturing the local degree of urbanization (thus based on population density) or the local labour market structure, based on commuter flows. This approach to human geography has the potential to portray a picture that is often different from that described by traditional administrative subdivisions.

1.3 Poverty analysis in the era of big data

The granular geospatial analysis of poverty and social phenomena hinges upon the availability of accurate and vast data on the living conditions of the local population. In the era of the "data revolution" (Einav and Levin, 2014a), this aspect can often be taken for granted, but in the quantitative study of social phenomena, it continues to represent a key hurdle, with this research being no exception. As previously discussed, the availability of new sources of microdata played a critical role in the development of the field of poverty analysis in the Italian context (Brandolini, 2021). While the Italian National Statistics Office (ISTAT) and the Central Bank had been collecting household surveys on consumption and income since the 1960s, it is under the initiative of the European Commission in the 1990s and 2000s that the production of large-scale micro-economic data for poverty analysis took a different turn (Jenkins, 2020). In 1995, the third European Community Poverty Action Program set, as its third main objective, the development of a statistical monitoring system (CEE (Commission of the European Communities), 1995). Under the aegis of EUROSTAT, a standardised "at risk of poverty measure" was introduced and, for the first time, individual- and household-level data, instead of tabulations by expenditure classes, formed the basis for its estimation (Hagenaars et al., 1994). This was possible following the launch of a harmonised European survey, the European Community Household Panel (ECHP), in 2004 replaced by the European Statistics on Income and Living Conditions (EU-SILC). The ECHP, conducted from 1994 to 2001, was the first detailed household income survey carried out by ISTAT and together with the Household Budget Survey, which replaced the Household Consumption Survey in 2009, has formed the basis for research on poverty studies over the last three decades. While maintaining a key role in the research and production of poverty statistics, these data sources present some inherent drawbacks. Firstly the complexity of the process of data collection and quality assurance means the reference period captured by the data pre-dates the publication schedule by almost two years, meaning the surveys become available with an inherent lag. Secondly, due to time constraints and financial limitations, the survey samples are designed to be representative at the national or macro area levels at best.

We are therefore some way off the promise of granular and near real-time household data that some thought the expansion of big data - here intended as the large amount of information generated by all sorts of digital human interactions - would make available. Overall, the use of big data in social science and public policy research remains limited (El-Taliawi et al., 2021) and some of the main concerns raised in the past about the application of these data sources in this field (White and Breckenridge, 2014) are still valid today. While providing information on many aspects of human interactions which was previously uncodified, indeed, there remain concerns about the quality of big data, which often results difficult to verify (Pratesi, 2017). Furthermore, big data are often collected from a self-selected group of individuals who are more likely to engage with technology or participate in online activities and can therefore lead to selection bias, undermining the generalizability of findings. This concern is especially valid in the field of poverty studies, where the population of interest is more likely to be systematically excluded from this type of digital interaction due to their socio-economic conditions. As a result, big data may not be representative of the entire population, and may only capture a limited scope of behaviours or experiences. However, if we expand the definition of big data beyond the type of information generated in the context of human interaction with digital services typically provided by private sectors actors, such as social media platforms, e-commerce or geo-localisation services, and we include, for example, publicly owned administrative data, then the disruptive potential of this data source in the social science research field remains vast (Connelly et al., 2016).

Administrative data refer to data that are collected and used for administrative purposes, such as by government agencies or other organizations, e.g., tax records, health care claims, educational records, and criminal justice data. Unlike traditional surveys, administrative data are usually based on larger samples allowing for analysis at more granular geographical levels, and, depending on the purpose for which is collected, it can be considered as representative of the entire population of interest (Meyer et al., 2015). Administrative data are often available in longitudinal format, making it particularly useful for studying changes over time, tracking the effects of public interventions, and making causal claims. The information collected is often verified by public officials with standardised procedures in order to minimize fraud and error in public service provision and this reduces the likelihood of errors or biases in the data. Finally, administrative data are often already collected and stored by government agencies or other organizations, which can reduce the costs of data collection and processing for researchers. In the last decade, National Statistics Offices and Government Agencies have increasingly relied on the use of administrative records for both research and operational purposes. In particular, great emphasis has been given to the integration of different administrative data sources as well as traditional surveys (Einav and Levin, 2014b). This approach not only augments the granularity of the data, but it improves their quality and allows to measure it. It also expands the information capacity of the data obtained by integrating different subject areas (individuals, households, economic units, and territory) through an integration that is not only physical but also conceptual.

In the Italian context, for example, in 2017, the first prototype of the Integrated System of Registers (SIR) was launched by ISTAT, as the basic infrastructure underpinning the new statistical information production model. The system aims to integrate data derived from administrative records, statistical surveys and new sources and ensure unified management of different themes (e.g., social, environmental or economic statistics) on the basis of a conceptual and physical integration between different statistical units (ISTAT, 2022b). In this framework, ISTAT reported that the direct collection of data from households, economic units and institutions within the National Statistical System between 2013 and 2019 decreased by 17.6% while the use of administrative sources increased by 33.5% (Alleva, 2018). Yet, despite this trend, access to administrative microdata for social science researchers remains limited due to confidentiality concerns, technical barriers and, in certain cases, bureaucratic hurdles. Frequently, as in the case of the research presented herewith, administrative records are made available exclusively at the aggregate level limiting the potential of the analysis. The finding of this work, for example, is based on the combination of traditional household surveys and administrative records. In particular, EUSILC and HBS surveys are the main sources of information to derive poverty indicators, complemented by a set of administrative registers related to tax returns, companies and enterprises registers, data on beneficiaries of public welfare programs, and voting data. The availability of these sources at the aggregate level dictated the choices of the estimation models employed in the analysis and determined the insight of the findings. Access to administrative microdata would have no doubt increased the relevance of the analysis. To conclude, therefore, as statistical systems are adapting to integrate administrative data as the main source of official statistics, so should new frameworks for granting secure access to this information to researchers in its microdata format.

Chapter 2

In-work poverty and sectoral employment. An analysis of local dynamics

Abstract: In-work poverty has risen to become a key feature of European societies over the course of the last decades. In 2017, the percentage of the working population at risk of low pay in Italy reached an estimated 25%. Yet, due to data limitations, few studies have analysed the local distribution of this phenomenon at the sub-regional level and have attempted to study the macro-determinant factors associated with its rise. By applying Small Area Estimates (SAE) to EU-SILC data we obtain a novel map of the geographical distribution of in-work poverty in Italy, defined as the share of workers at risk of low pay (AROLP), for the period between 2008 and 2017. The unit of analysis of Local Labour System areas, a non-administrative geographical unit based on commuter flows, highlights the deepening of Italian dualism between Northern and Southern areas, which overwhelmingly reported the largest increases in AROLP rates, as well as rising within-region wage inequality. By matching the small area estimates for AROLP with data on local sectoral employment composition and other employment indicators, we study how dynamics of growth and decline in specific sectors are associated with trends of in-work poverty, in relation to the process of de-industrialisation. By means of a panel fixed effect regression model, we observe that growth in low-skill and low-productivity sectors such as agriculture and commerce are associated with increases in AROLP incidence. On the contrary trends of low pay are negatively associated with the growth of manufacturing jobs, admin and support services to enterprises, and technical and scientific professions. In addition, variations in overall employment rep-

resent the strongest predictor for dynamics of low-pay incidence. The analysis reveals trends at odds with the literature on sectoral employment and wages in relation to the sectors of hospitality and health and social care. Driven by areas in southern regions, the growth of these sectors in Italy is associated with lower levels of low-pay incidence.

2.1 Introduction

A key feature of modern poverty regimes across industrialised economies is the increasing presence of individuals in employment among those identified as poor (Edmiston, 2022; Saraceno and Benassi, 2020). The combination of structural changes related to processes of de-industrialization as described in Section 1.2 and the shift towards workfare paradigms of social security is commonly seen as the drivers behind this trend (Snel et al., 2008; Sawicky, 2002; Lohmann, 2008). As a result, over the course of the last decades, developed economies witnessed a rise in in-work poverty, with Italy being no exception. An independent report commissioned by the Italian Ministry of Labour has estimated that, in 2017, the percentage of households considered at risk of low pay reached 25% of the working population, rising from 18% in 2008 (Garnero et al., 2021). The reduction in unemployment that has taken place in many European countries during the same period, following the economic crash of 2008, indeed, has not meant a substantial reduction in the phenomenon of poverty. This is related, in part, to the emergence of a segment of the population that, while working, receives an income below the poverty line: the so-called working poor. The growth of in-work poverty demonstrates that while employment can be an important route out of poverty, the quality of jobs is critical in providing a sustainable exit from poverty (Tomlinson and Walker, 2010). Despite the rising prominence of the phenomenon and the increasing attention by social scientists in the analysis of these trends, work poverty is a relatively recent emerging topic in the literature on social policy and labour economics. As a result, studies focusing on the issue have devoted little focus to the sub-national distribution of in-work poverty. The origin of this gap in the literature is directly related to data availability, as most studies on in-work poverty are based on national surveys. This type of survey is designed to provide reliable indicators at the national or, at the most, regional level, but due to the limited samples, can not be employed for more granular geographical analysis. Furthermore, detailed administrative data not easily accessible is often required to obtain such local estimates. By capturing the geographical heterogeneity, sub-national analyses not only may provide a more accurate description of the phenomenon, but they would also allow us to explore its association with local determinants.

Furthermore, most studies to date have focused on evidence related to the microdeterminants of in-work poverty, identifying individual characteristics that negatively affect the probability of accessing favourable labour market arrangements (e.g., young age and low level of education) or characteristics and structure of the household (e.g., having a small number of earners and a high number of dependent children)(Barbieri et al., 2018). Filandri and Struffolino (2019) provides an important assessment of the macrodeterimants of in-work poverty related to job market characteristics, but focuses exclusively on female workers. Sissons et al. (2018) investigate the probability of households living in poverty based on the employment sectors of household earners. They find that working in so-called low-productivity sectors, such as hospitality and transportation increases the probability of living in poverty. Their study, however, provides only a static picture of the link between sectoral employment and low pay thus falling short of describing how the dynamics related to the growth and decline of specific sectors are associated with the rise of in-work poverty. This chapter tries to address these gaps in two ways. First, it applies Small Area Estimation methods to provide a reliable geographic breakdown of the phenomenon of in-work poverty in Italy at a finer geographical level than provinces (NUTS3) and without access to administrative micro-data on employment, describing how the phenomenon evolved over the period between 2008 and 2017. Secondly, it explores how variations of in-work poverty at the local level are associated with shifts in sectoral employment based on administrative registers of economic activities. Sectors are a common way of dividing types of employment and are characterised by very different conditions and average skill levels, aspects strictly related to dynamics of low pay. In addition, they represent a useful unit of analysis for interpreting results in light of sector-level projections about those parts of the economy set to grow and those set to decline.

The units of analysis adopted hereby are the so-called Local Labour Systems (LLS) areas, 610 non-administrative geographical units whose boundaries are drawn by ISTAT based on commuter flows. The very nature of this unit of observation, defined on the basis of local employment characteristics and not on administrative boundaries, aptly fits the focus of the analysis. As discussed in Section 1.2, non-administrative spatial units can help reveal geographical trends often concealed by mapping based on traditional administrative boundaries. The first contribution of this research is to provide, to the best of our knowledge, the first account of in-work poverty at the sub-regional level in Italy and beyond. Secondly, by focusing on the growth and decline of economic sectors and their association with trends of low-pay incidence, it provides fresh evidence to the literature on the geography of deindustrialisation and its impact on society. The results and conclusions drawn by our research, therefore, provide novel evidence for researchers and policymakers alike on the geographical composition of the phenomenon of in-work poverty, while casting light on the dynamics that link it with macroeconomic trends related to employment. The structure of the paper is the following. Section 2.2 presents the key definitions for our analysis and background of the context and literature related to the phenomenon under study, Section 4.4 presents data used and some descriptive statistics. Section 4.4 hosts SAE models and results, before discussing findings related to the sectoral employment correlation analysis in Section 2.5 and concluding in Section 4.7.

2.2 Definitions and background

2.2.1 At-risk-of-low pay

The literature provides two definitions of the working poor, which are analytically distinct but empirically strictly related. The first definition, employed by EUROSTAT, considers the working poor as those individuals who are in work and live in a household with income below the poverty line. This definition thus combines the employment status of the individual and the household (equivalent) income, which identifies the worker's poverty status. The focus is not on individuals' remunerations in the labour market but rather on total household income, departing from the assumption of perfect pooling of resources within the household. While this assumption is usually adopted in distributional analyses to measure individual economic well-being, a number of empirical studies have shown how theories of non-unitary household behaviour more accurately reflect the decision-making dynamics over the allocation of resources within a household (Duflo, 2003; Bennett, 2013). In these models, decisions over the allocation of consumption are taken by negotiating partners whose bargaining power depends on the resources they command when the relationship breaks down (Lundberg and Pollak, 1996).

The second definition, and the one utilised in this research, considers working poor those individuals whose earnings are below the 60% of median earnings. Also referred to as at-risk-of-low-pay (AROLP), it is solely centred on the individual dimension with no reference to the household, and it identifies the working poor as low-paid workers. As a result, earnings are the key mechanism behind workers' economic circumstances. This definition establishes a direct link between labour market characteristics and the phenomenon of in-work poverty, and better serves the purpose of the study exploring the link with shifts in local sectoral employment. Such an approach, however, often overlooks the impact played by low work intensity and by discontinuous employment on low wages (Filandri and Struffolino, 2019). For the purpose of this research, we expand this definition by considering three different indicators of at-risk-of-low pay, based on hourly, weekly and annual earnings. These indicators, combined together, incrementally capture those factors related to salary structures and labour market dynamics which determine the incidence of low pay. A low-pay indicator based exclusively on hourly earnings describes the phenomenon of low pay as solely the result of low unit wages. The weekly earnings indicators add to the unitary wage dimension the one related to work intensity. Finally, the annual wage definition adds the dimension of discontinuous employment as the result of short-term temporary work contracts and seasonality.

2.2.2 Background

As discussed in Chapter 1.1, the period of observation considered in this study, spanning from the year 2008 to 2017, marked a phase of deep economic transformation for the Italian economy. The effect of the Great Recession and subsequent sovereign debt crisis had severe repercussions on the country's employment levels, while at the same time accelerating the process of de-industrialization and contraction of the Italian manufacturing sector which had started a decade earlier. In this context, as previously mentioned, average real wages in 2017 were 1% lower than in 2020, a trend largely driven by lower nominal pay increases at the bottom of the wage distribution resulting in increasing wage inequality.

No agreement exists among economists on the drivers behind the trend of wage stagnation and rising wage inequality as recorded by industrialised economies over the course of the last decades. While classic economic models were built on the assumption of a strong link between labour productivity and compensations, a growing number of authors point to the gap between growth in productivity and wages observed in the last four decades as evidence of other factors determining wage levels (Bivens and Mishel, 2015; Schwellnus et al., 2017). In particular, the decline in the share of the workforce represented by trade unions (Lawrence, 2016; Machin, 2016) and public policy decisions around employment law and the setting of a minimum wage are seen as the primary factors behind the wage-productivity gap (Bivens et al., 2014). On the contrary, supporters of market solutions continue to argue that the link between wages and productivity remains strong (Strain, 2019). Stansbury and Summers (2018) provide substantial evidence of the linkage between productivity and compensation, whether this is measured as workers' median and average wage or as compensation among production workers. Lawrence (2016) demonstrates how the argument of the wage-productivity gap is heavily sensitive to the way the two indicators are measured, with his proposed approach showing similar trends in compensation and productivity. Whether the result of sluggish productivity or low employment protection and trade union representation, empirical evidence highlights how the process of de-industrialisation, described as the decline of the labour force employed in the manufacturing sector, is associated with stagnating wages at the bottom of the income distribution, wage inequality (David and Dorn, 2013) and in-work poverty (Cormier and Craypo, 2000). There is indeed increasing evidence that the wage-productivity gap is higher in the service sector as compared to manufacturing (Berlingieri et al., 2017), where the gains from productivity growth are transferred into wages with a higher elasticity (Hirschman, 1958; Szirmai, 2012). The main drivers behind this trend are identified in the lower presence of trade union members among service service sectors workers (Coveri and Pianta, 2022) and the diffuse practice of atypical employment types within these industries (Kaduk et al., 2019). Reduced wage-setting bargaining power and higher recourse to outsourcing and discontinuous employment are indeed associated with lower wages (Dube and Kaplan, 2010). As a results, a study of cross sector wage differential in the United States found that manufacturing workers earn 13.0% more in hourly compensation (wages and benefits) than comparable workers earn in the rest of the private sector (Mishel, 2018). On the contrary, in the UK context, Sissons et al. (2018) finds a positive and significant correlation between being employed in so-called low-productivity service sectors, such as hospitality and transportation, and the probability of living in poverty.

2.3 Data and descriptive statistics

2.3.1 EU-SILC survey

Estimates for the at-risk-of-low pay indicator (AROLP) across all 610 local labour systems (LLSs) are based on European Union Statistics on Income and Living Conditions (EU-SILC) survey data collected in the years 2008, 2015, 2016 and 2017. The choice of the period of observation is motivated by the unique macroeconomic conditions described in Section 2.2. Due to data availability limitations, however, it was not possible to obtain survey data containing the required level of geographical detail for all the years between 2008 and 2017. EU-SILC aims at collecting timely and comparable cross-sectional and longitudinal multidimensional microdata on income, poverty, social exclusion and living conditions in Europe.

In addition, it reports detailed information on respondents' employment status, the number of weekly hours worked, the number of months worked within the reference year as well as hourly, monthly and yearly wages. While it does not constitute an administrative data source from tax returns or social security payments, it is routinely utilised by national statistics offices (EUROSTAT, 2022a) for analyses related to low work intensity and by researchers likewise for analysis on in-work poverty (Filandri and Struffolino, 2019).

For the sake of clarity, we present the findings of the analysis of EU-SILC data for the years 2008 and 2017 only. The results for the 2015 and 2016 waves are part of the panel data employed to explore the correlation with sectoral employment dynamics. Findings of the SAE models for these years are reported in Appendix 2.B. The 2008 and 2017 waves of the survey contain information on self-reported income for the years 2007 and 2016 with a total of 18,0000 and 22,200 observations respectively. As reported in Table 2.1 and in line with the findings of a recent study commissioned by the Italian Ministry of Labour (Garnero et al., 2021), EU-SILC data show a significant increase in the overall phenomenon of low-pay at the national level, which, if we consider the most expansive of definitions, reached almost a quarter of all workers in the year 2017.

Table 2.1: National percentages of AROLP based on annual, weekly and hourly earnings.

	Annual earnings	**Weekly earnings**	**Hourly Earnings**
EUSILC 2008	18.4%	16.6%	14.4%
EUSILC 2017	23.3%	22.6%	20.9%

EU-SILC has a two-stage sample design where strata are regions by type of municipality. The sampling structure of the EU-SILC survey, however, is constructed in a way that estimates at lower-geographical levels than regions (NUTS2) can not rely on large enough samples to be considered reliable. LLSs are unplanned domains that cut across sampling strata and provincial (NUTS3) areas. LLSs are sub-regional geographical areas where the bulk of the local labour force lives and works, and where employers can draw from in order to fill in the majority of vacancies. In Italy, there are 610 distinct and functional areas defined as clusters of municipalities through an allocation process based on commuting patterns collected by the 2011 Population Census. Table 3.3.1 reports the distribution of the sample sizes of the EU-SILC for the 610 LLSs considered in this analysis. As we can see, in both years the minimum number of sampled households is equal to only one respondent household. As a result 'direct' estimates - that is, estimates computed using only survey data and sampling weights - at this level of analysis are likely to be accompanied by high level of uncertainty.

Table 2.2: Sample size distribution of EU-SILC surveys across the 610 LLSs.

	Min	**1st Qu.**	**Median**	**3rd Qu.**	**Max**	**Out of Samp**
EU-SILC 2008	1.00	19.00	32.00	64.25	716.00	250
EU-SILC 2017	1.0	19.0	31.50	59.25	914.0	290

In this context, Statistics Canada (Wannell and Usalcas, 2012) provides guidelines for publication related to the uncertainty of estimates: estimates with a Coefficient of Variation (CV) less than 16.6% are considered reliable for general use; estimates with CVs between 16.6% and 33.3% should be accompanied by warnings to users; and estimates with coefficients of variation larger than 33.3% are considered unreliable. Table 2.3 shows the number of areas classified according to these three thresholds of the CV on the basis of the 'direct' estimates for the years 2008 and 2017. As we can see, out of the 610 total areas of interest, the number of areas with a CV in the second or third class of CV values is rather high for all indicators, suggesting the need to resort to appropriate modelling techniques to reduce the estimates' CVs.

Table 2.3: Coefficients of variation, out of sample units of ARLOP direct estimates based on hourly, weekly and annual earnings for the EUSILC waves 2008 and 2017.

2008	<16.5%	16.5-33.3%	>33.3%	NA	Out of sample
$AROLP_{hourly}$	2	65	271	22	250
$AROLP_{weekly}$	5	82	250	23	250
$AROLP_{annual}$	5	89	246	20	250
2017					
$AROLP_{hourly}$	9	87	209	15	290
$AROLP_{weekly}$	10	92	204	14	290
$AROLP_{annual}$	10	99	198	13	290

The basic idea of Small Area Estimation (SAE) techniques is to introduce a statistical model to exploit the relationship between the variable of interest and some covariates for which population information is available in order to improve the precision of direct estimates. We consider as auxiliary variables, a set of administrative covariates coming from the Italian Ministry of Treasure Tax returns data referred to years 2008, 2015, 2016 and 2017 at the municipality level. The variables employed in the models as auxiliary information are (i) the percentage of taxpayers, (ii) the percentage of the population with estates (iii) the percentage of the population with yearly income below €10,000, (iv) the percentage of the population with yearly income between €10,000 and €15,000 (v) the percentage of the population with yearly income between €15,000 and €26,000, (vi) the percentage of the population with yearly income between €26,000 and €55,000 (vii) the average estate and the business income.

2.3.2 ASIA registry and Cambridge Econometrics data

Finally, we rely on the ASIA registry collected by ISTAT to study the association between sectoral employment dynamics and variations in low-pay incidence at the local level. The registry consists of information on the nature and structure of economic units aggregated by the NACE 1-digit taxonomy at LLS level. The register does not include economic activities related to agriculture, forestry and fishing (NACE code "A") and public sector and non-market activities (NACE codes "O-U"). For these sectors, we rely on data from Cambridge Econometrics collected at the NUTS3 level and subsequently assigned to each LLS based on its population share within each NUTS3 area. To obtain a measure of the intensity of local sectoral employment we divide the total number of workers in each NACE 1-digit sector by the active population within each LLS. We then focus only on those sectors with an average share of the employed active population of 2.5% or above, as reported in Table 2.4. As illustrated in Figure 2.1, Figure 2.2 and Figure 2.3, we observed considerable variations in the local employment make-up across sectors during the reference years 2007 and 2008. Sectors such as manufacturing and construction, for example, display a significant decrease in the overall share of employment of the local active population across all three main geographical areas of the country. Other sectors, such as commerce, transport and logistics, hospitality professions and technical activities, and rentals and support services display a less clear trend with significant heterogeneity at the local level. Health care and social service and non-market services are the only sector clearly displaying an increase. Employment, as previously described in Section 2.2, presents an overall decrease, with areas in the South showing the sharpest drop. These findings are closely aligned with evidence on sectoral employment trends during the same period in other European economies (Foster-McGregor et al., 2012).

Table 2.4: Employment variables descriptive statistics

	Mean	SD	Min	Max
Manufacturing	0.175	0.103	.018	0.641
Constructions	0.081	0.018	0.021	0.184
Wholesale and retail trade	0.144	.025	.061	0.377
Transporting and storage	0.046	.020	.007	0.139
Hospitality	0.048	0.033	.008	0.889
Professional and technical activities	0.049	0.019	0.012	0.098
Administrative and support activities	0.043	0.023	0.001	0.109
Health and social care	0.026	0.009	0.004	0.078
Agriculture	0.042	0.037	0.004	0.173
Public sector	0.284	0.063	0.128	0.415
Payroll	0.507	0.038	0.280	0.687
Selfemployed	0.019	0.006	0.002	0.033

Notes: Number of observations 610, all variables expressed as share of active population.

Figure 2.1: Variations in the proportion of the workforce employed by sector at LLS level. Sectors with >2% of local workforce only. In red are represented LLSs in the north of Italy, in black LLSs in the centre of Italy, in green LLSs in the south of Italy. Black line represents the 45° bisector.

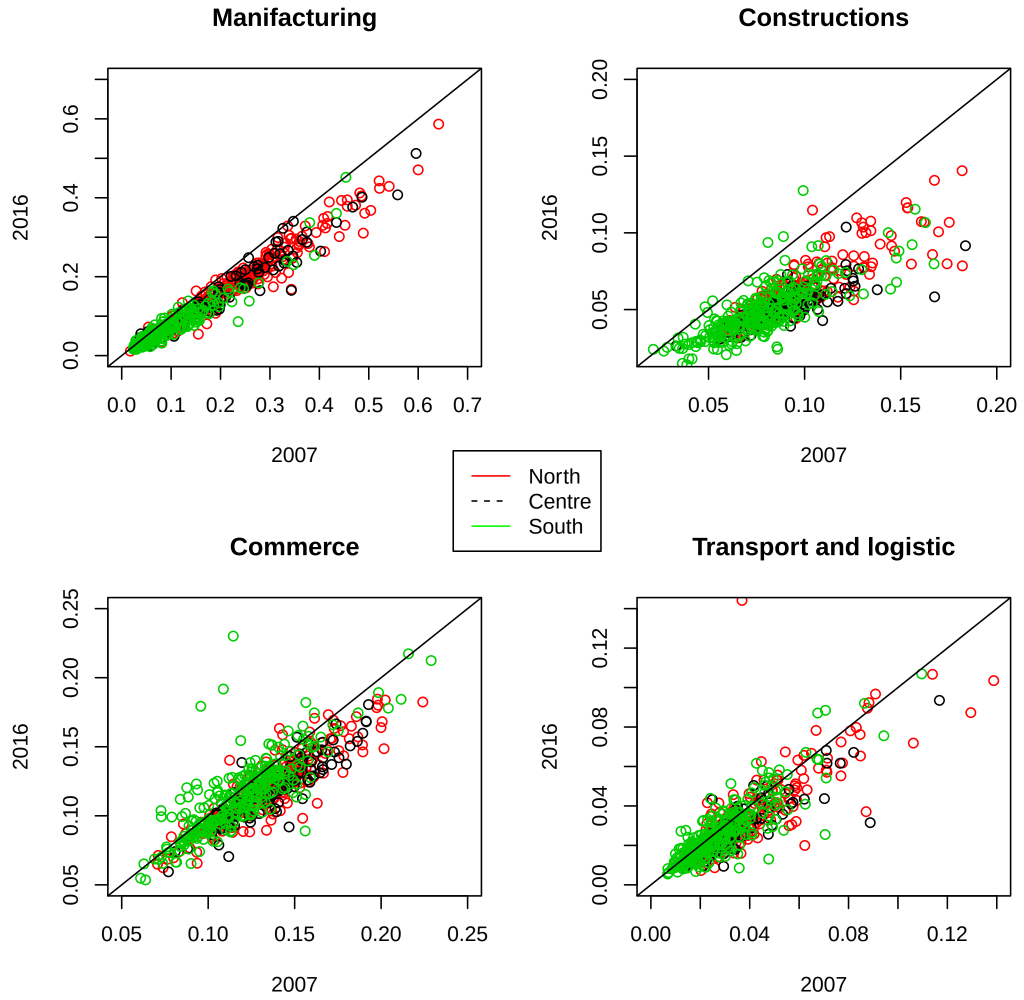


Figure 2.2: Variations in the proportion of the workforce employed by sector at LLS level. Sectors with >2% of local workforce only. In red are represented LLSs in the north of Italy, in black LLSs in the centre of Italy, in green LLSs in the south of Italy. Black line represents the 45° bisector.

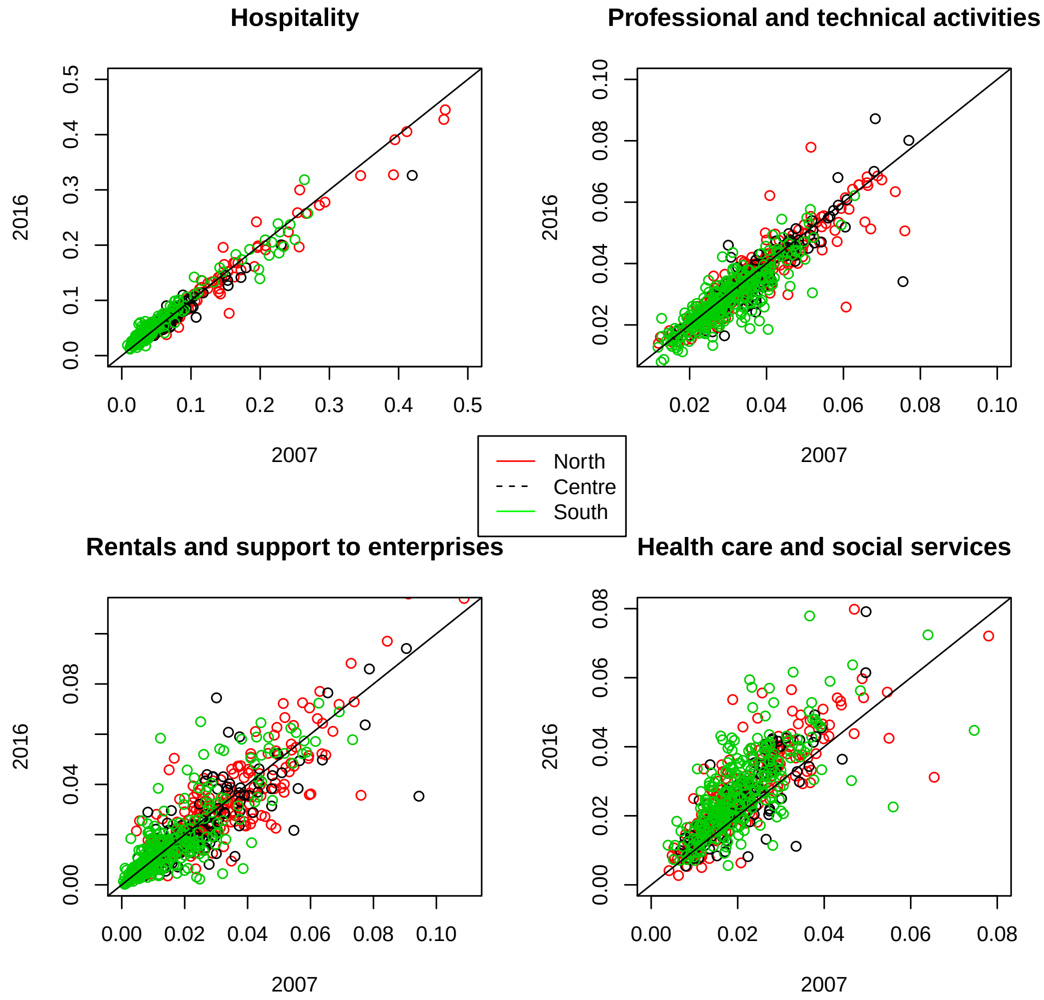
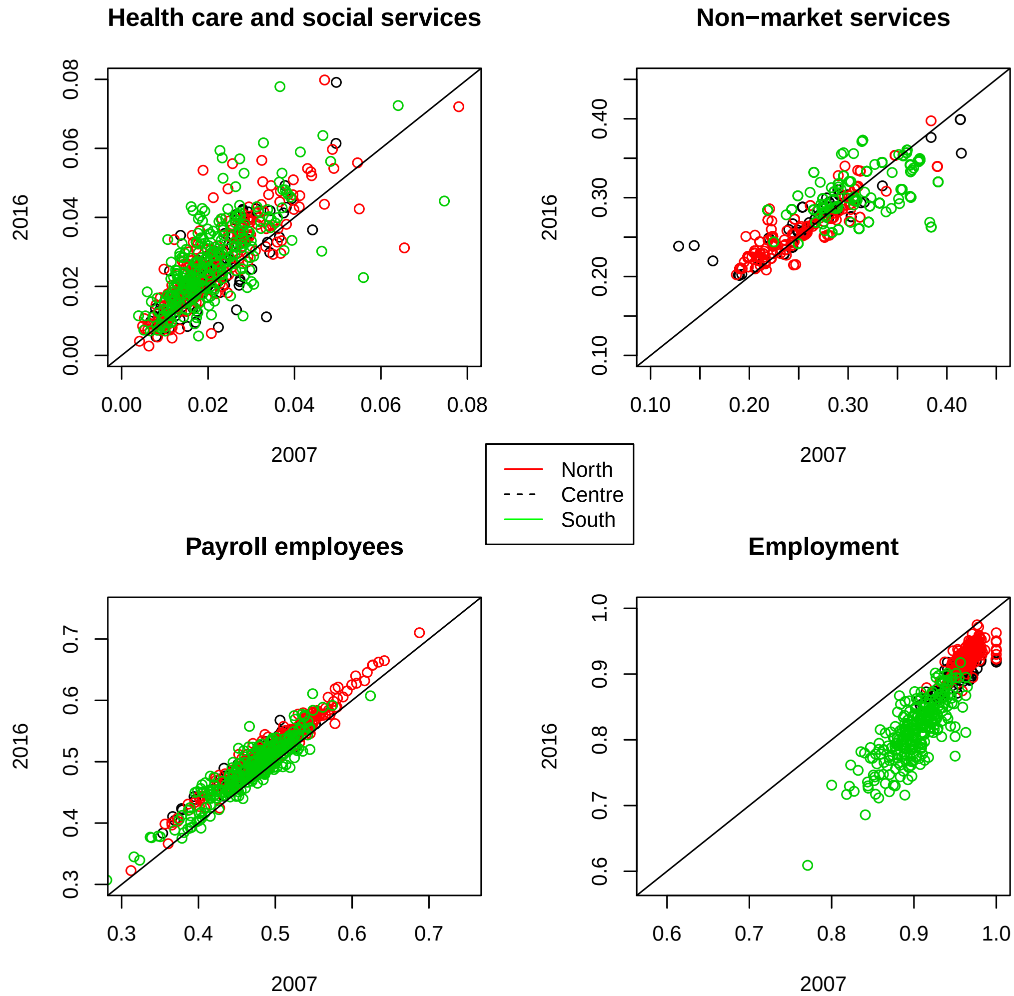


Figure 2.3: Variations in the proportion of the workforce employed by sector at LLS level. Sectors with >2% of local workforce only. In red are represented LLSs in the north of Italy, in black LLSs in the centre of Italy, in green LLSs in the south of Italy. Black line represents the 45° bisector.



2.4 Small Area Estimates

2.4.1 Small Area Estimation models

In this section, we describe the methods employed for obtaining estimates at the LLS level for the three indicators of at-risk-of-low pay described hitherto. For all three definitions, our target indicators are the small area proportions. The application of SAE models aims at increasing the precision of direct survey estimates through the use of the administrative covariates at LLS level. In this application, we make use of the Fay and Herriot (FH) model (Fay III and Herriot, 1979) with an arcsine-transformation of the direct estimates Casas-Cordero Valencia et al. (2016), Burgard et al. (2015), and Schmid et al. (2017). The FH model and its transformations are area-level models that link direct estimates to area-level covariates, for the estimation of the ARLOP indicator. They are especially useful when access to individual-level data is not available, as the auxiliary variables and the direct estimators only need to be available on an aggregated level.

Arc-sine transformed Fay-Harriot model

The area level model proposed by Fay and Herriot (1979) (the FH model) links the direct estimates obtained from survey observations with synthetic area level estimates obtained through error-free covariates. The FH model is based on two stages. Let us assume that there are m small areas of interest and that θ_i represents the population characteristic of interest in area i, such as a mean, with $i = 1, \ldots, m$. The sampling model (first stage), links the θ_i to their direct estimates $\hat{\theta}_i^{direct}$ provided by a survey as follows:

$$\hat{\theta}_i^{direct} = \theta_i + e_i \tag{2.1}$$

where $e_i \sim N(0, \sigma_{ei}^2)$, $i = 1, \ldots, m$ are the independent design errors. The second stage model relies on a set of known covariates x_i for each area i as such:

$$\theta_i = x_i^t \beta + z_i v_i \tag{2.2}$$

where β is the vector of model parameters, z_i are known positive constants and $v_i \sim N(0, \sigma_u^2)$ are independent and identically distributed random area effects, with u_j independent from e_i for all i and j. The combination of both models leads to an area-level linear mixed model given by:

$$\hat{\theta}_i^{direct} = x_i^T \beta + z_i v_i + e_i \quad (2.3)$$

This model relates direct estimates to specific domain covariates, considering the random area effects as independent. The target parameter is referred to as a composite estimate. Indicators, such as the ARLOP considered in this paper, are expressed within an interval range of [0,1]. FH models however offer no guarantee to provide estimates within such range. Transforming the target parameter to approximate a normal distribution of the error terms for a better fit of the model covariates has been done in multiple studies. Here we consider an inverse sine transformation $h(x) = sin^{-1}(\sqrt{x})$ as in Casas-Cordero Valencia et al. (2016), Burgard et al. (2015), and Schmid et al. (2017). Hadam et al. (2020) propose a bias-corrected back transformation which allows for the analytical solution for the estimates of MSE and confidence intervals through parametric bootstrap. Following Jiang et al. (2001), the model considers the variance for the transformed direct estimates $(sin^{-1}(\sqrt{\hat{\theta}_i^{direct}}))$ using the effective sample size n_i, thus obtaining $\sigma_{ei}^2 = 1/4n_i$.

Assuming the independence and normality of error terms and area random effects, the model is specified as follows.

$$(sin^{-1}(\sqrt{\hat{\theta}_i^{direct}})) = x_i^T \beta + v_i + e_i, \quad (2.4)$$

As described above, the parameters β and u_i can be estimated, leading to the FH estimator at the transformed level.

$$\theta_i^{FH} = \gamma_i(sin^{-1}(\sqrt{\hat{\theta}_i^{direct}})) + (1 - \gamma_i)x_i^T \hat{\beta}, \quad (2.5)$$

where $\hat{\gamma}_i = \frac{\hat{\sigma}_v^2}{\hat{\sigma}_v^2 + \sigma_{ei}^2}$ is defined as the shrinkage factor, assuming values in range $[0, 1]$. It, therefore, assigns a weight to the indirect estimate θ_i and by definition assumes values $\gamma_i = 0$ for any out-of-sample area. The predictor $\hat{\theta}_i^{FH}$ is therefore a convex combination of a transformation of the direct estimator $\hat{\theta}_i^{Direct}$ and the predicted value $x_i^T \hat{\beta}$ from the regression model.

In order to obtain the target parameter, is, therefore, necessary to operate a back-transformation of the FH estimates. Jensen et al. (1906) have demonstrated how the naive back-transformation suffers from bias due to the nonlinearity of the transformation. Hadam et al. (2020) propose a solution for correcting for this bias, using numeric integration and evaluating their results against official administrative estimates of the target parameter.

While previous applications of this transformed FH model provided an analytical solution to the estimation of confidence intervals only (Casas-Cordero Valencia et al., 2016; Schmid et al., 2017), Hadam et al. (2020) propose a novel approach for the estimation of the MSE. The approach follows González-Manteiga et al. (2008) bootstrap procedure, by including the bias-corrected back-transformed FH estimates.

Benchmarked Fay-Herriot estimators

The model presented above provides estimates for all 610 LLS on the Italian territory. However, the aggregated estimates at the national level can differ from the corresponding direct estimator. According to the theory of small area estimation, the parameters β and σ_v^2 are unknown and must be estimated while the σ_{ei}^2 are assumed to be known. The estimators of the σ_{ei}^2 are often smoothed, and the smoothed estimators are treated as if they were the true sampling variances (Rao and Molina, 2015).

Following Datta et al. (2011) we apply a benchmark approach to achieve internal consistency with the direct estimator both at the nationa and regional level (regional benchmark are obtained via direct estimates, as reported in Appendix 2.A) so that:

$$\sum_{i=1}^{D} \xi_i \hat{\theta}_i^{FH,bench} = \tau,$$

where ξ_i represents the share of the population size of each area over the total population size (Ni/N). In our application, the EBLUP estimators are aggregated at the regional level and then at the national level. The benchmarked FH estimator is defined by Datta et al. (2011) as:

$$\hat{\theta}_i^{FH,bench} = \hat{\theta}_i^{FH} + (\sum_{i=1}^{D} \frac{\xi_i^2}{\phi_i})^{-1}(\tau - \sum_{i=1}^{D} \xi_i \hat{\theta}_i^{FH}) \frac{\xi_i}{\phi_i}$$

There are several ways to define the weight ϕ_i. For both FH estimators proposed in the previous sections, we use a naive approach where the weights are given by $\phi_i = \xi_i$.

2.4.2 Results: SAE estimates at LLS level

Small area models are employed to improve the precision of direct estimates of the three AROLP indicators from the EU-SILC waves considered in this study. To facilitate the reading of the paper, we report the results for the waves 2008 and 2017, which define our observation period. The maps and the CVs of the estimates for the other years are reported in Figure 2.B.1, Figure 2.B.2 and Table 2.B.1 in Appendix 2.B. To assess gains in the accuracy of our estimates we compare the coefficient of variations of arcsine transformed FH model with those of the respective direct estimates.

In this analysis, the application of SAE methods brings considerable gains to the precision of estimates as illustrated in Table 3.5.1 compared to Table 2.3. The arcsine transformed FH model brings almost all CV estimates below the 33.3% threshold for all indicators across both years, with only three and four local LLS areas with a precision of estimates too low to be considered reliable for estimates from the EUSILC 2017 wave. In addition, the p-values of the Brown test for all six models reject the null hypothesis that FH and Direct estimates are statistically different, evidence of the goodness of fit of the models.

The findings of the analysis at the LLS area level highlight a stark widening in the gap between the North and the rest of the country with respect to the incidence of the phenomenon of AROLP, as presented in Figure 2.4, Figure 2.5 and Figure 2.6. Areas in the regions of Sardinia, Abruzzo, Sicily and Calabria show the highest increase in incidence with increases of up to 31% as illustrated in Figure 2.7. On the contrary, 102 out of 123 areas which recorded a decrease in AROLP incidence as per one of the three definitions considered so far are located in Northern regions. A comparison of the estimates across the three definitions of AROLP incidence highlights how once hourly earnings are considered, the regional divide is somehow smaller. An analysis of variations by macro areas, indeed, indicates how the hourly wage indicator follows a more homogeneous trend, with regions in the North recording on average an increase of 4.1pp. against an 8.0pp increase among Southern regions. This gap between North and South is significantly smaller than those recorded for the annual and weekly earnings indicators, where we observe an increase of 2.0pp and 2.7pp respectively in Northern regions and 9.1pp and 10.9pp in the South. These findings suggest how low work intensity and discontinuous employment, as opposed to unit-wage disparity, seem to be more strongly associated with the increasing geographical wage inequality across Italy's main geographical areas.

Considering the LLS as the main unit of analysis allows us to uncover important heterogeneity within bordering areas and areas located within the same regional borders. The country's central area displays the strongest heterogeneity with LLSs in regions such as Tuscany, Umbria and Marche displaying values of AROLP incidence raging across three quartiles of the overall distribution (see panel b in Figure 2.4, Figure 2.5 and Figure 2.6). When we shift our focus onto the variations in AROLP incidence we notice how such heterogeneity expands to other areas and regions too, especially in the South, where clusters of areas with high increases in AROLP are dotted by units recording much more moderate variations. Figure 2.7, furthermore, reveals the presence of clusters of areas displaying similar trends across different regions.

Table 2.5: Coefficients of variation, Brown test p values and correlation of at-risk-of low pay estimates (annual earnings), obtained with FH arcsine models. The estimates were benchmarked at the regional level using the direct estimates for the region as benchmark.

2008	<16.5%	16.5-33.3%	>33.3%	Brown test	Corr. Direct
$AROLP_{annual}$	593	17	0	p=0.991	0.673
$AROLP_{weekly}$	522	88	0	p=0.999	0.673
$AROLP_{hourly}$	210	400	0	p=0.999	0626
2017					
$AROLP_{annual}$	532	75	3	p=0.06	0.681
$AROLP_{weekly}$	544	63	3	p=0.13	0.685
$AROLP_{hourly}$	339	267	4	p=0.979	0.721

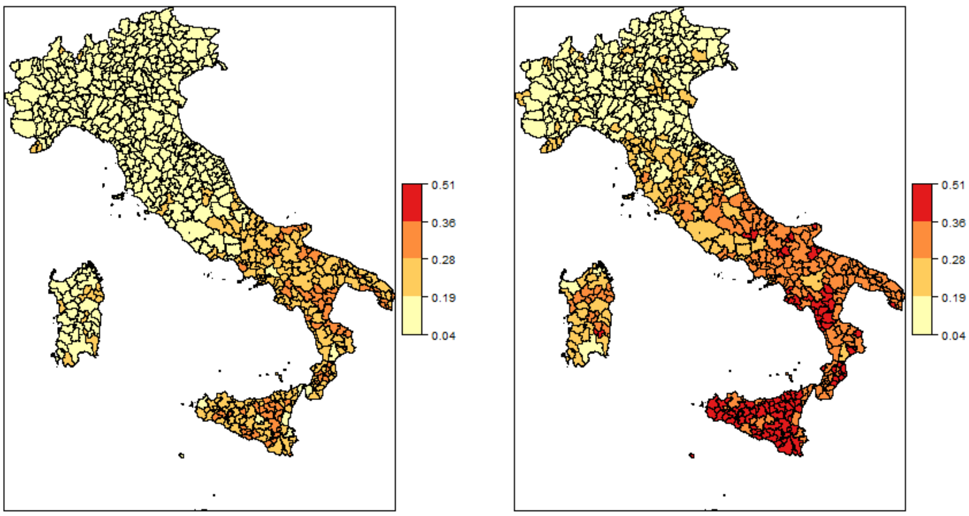


Figure 2.4: Arc-sin transformed FH estimates for AROLP based on hourly income by LLS for the year 2008 (left panel) and year 2017 (right panel)

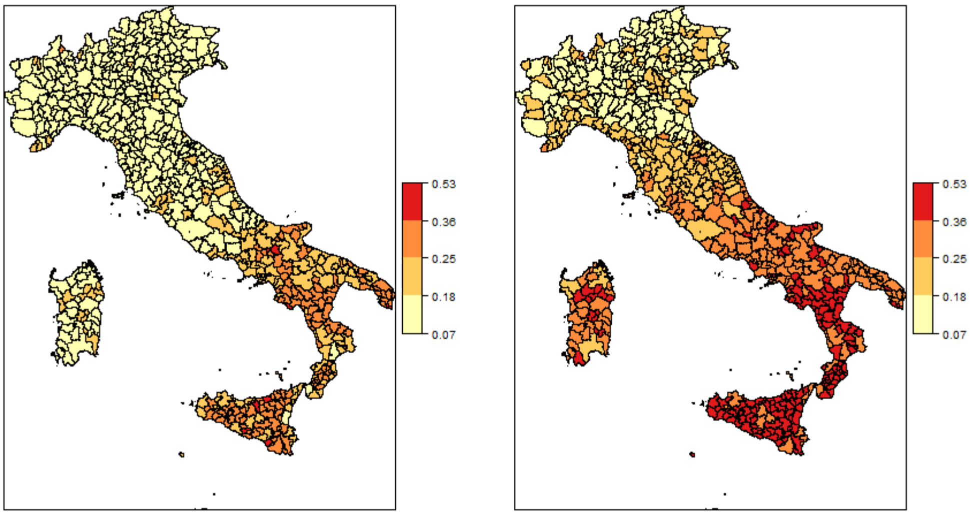


Figure 2.5: Arc-sin transformed FH estimates for AROLP based on weekly income by LLS for the year 2008 (left panel) and year 2017 (right panel)

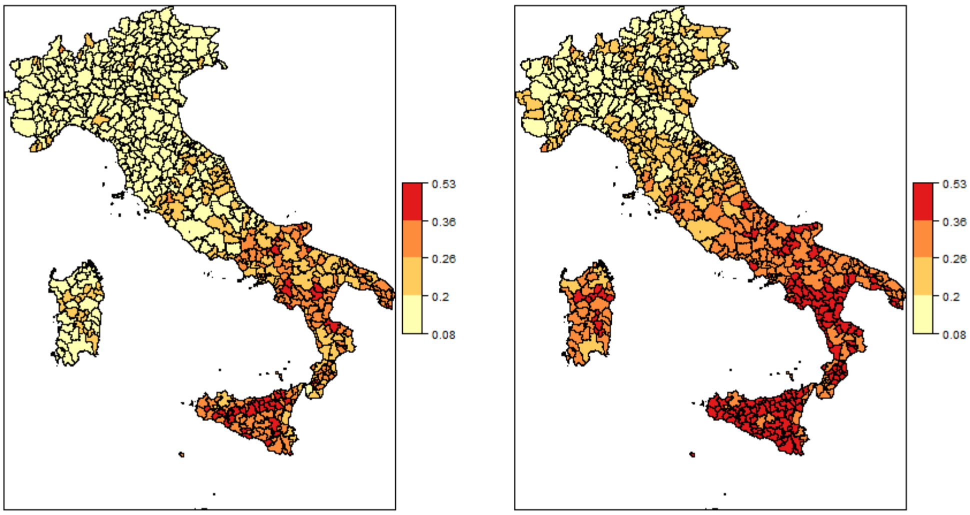


Figure 2.6: Arc-sin transformed FH estimates for AROLP based on annual income by LLS for the year 2008 (left panel) and year 2017 (right panel)

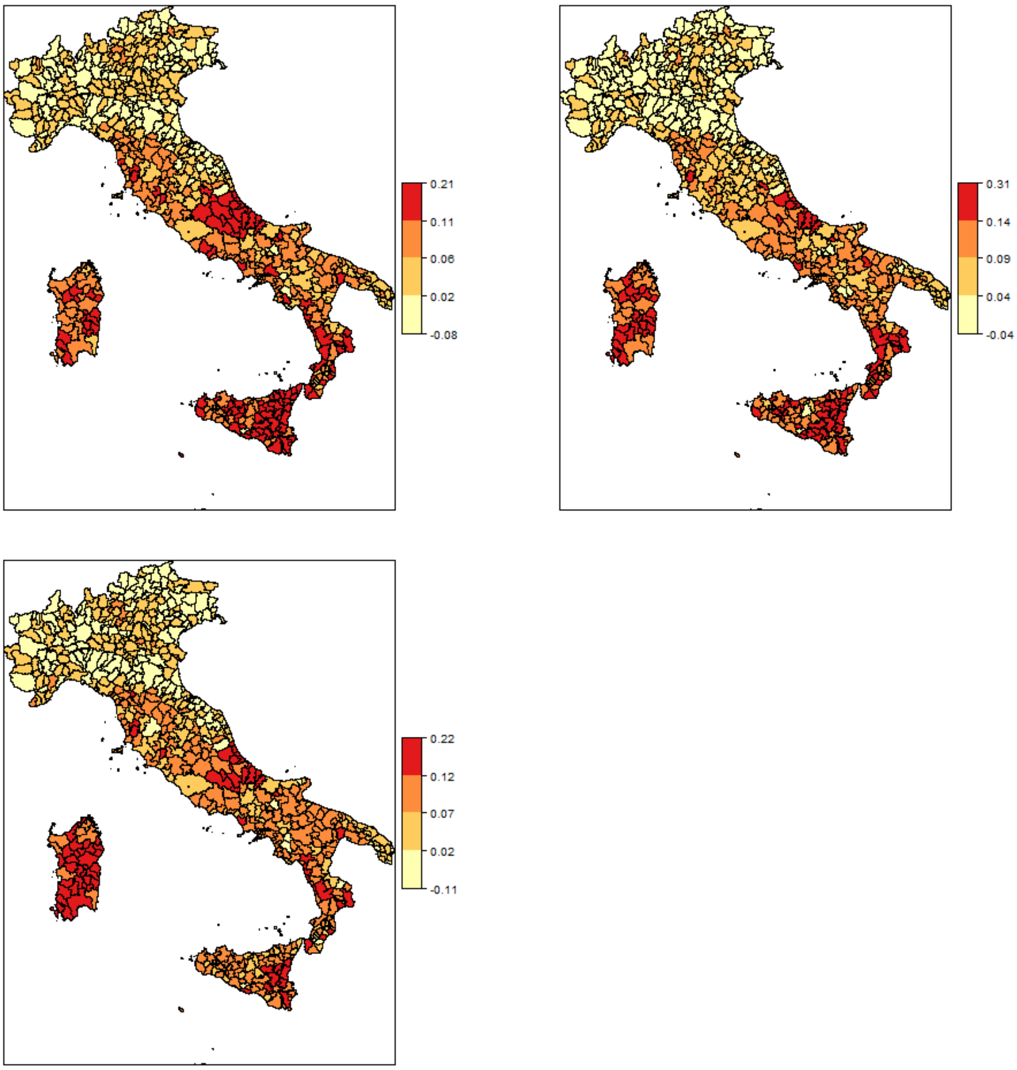


Figure 2.7: Changes in at-risk-of low pay rates by LLS between 2008 and 2017 based on annual income (panel a), weekly income (panel b) and hourly income (panel c)

An analysis of the LLS estimates by region confirms the strong trend in the geographical distribution of the phenomenon of low pay, marking a net difference between northern, central and southern areas. As reported in Figure 2.8, 2.9 and 2.10, no region has witnessed a decrease in the AROLP incidence, regardless of which indicator we consider. However, the overall increase recorded at the national level appears to be driven by the centre and southern regions. Across all three definitions of AROLP indicators, the five regions recording the highest increase in the phenomenon of low pay, indeed, are all located in the centre and southern part of the country, with Abruzzo, Calabria and Sicily consistently featuring within this group. Interestingly, the analysis at the LLS level allows us to unmask important heterogeneity within the same region. As illustrated in Figure 2.8, 2.9 and 2.10 the variability within regions increases significantly in 2017, across the majority of regions, irrespective of the broad macro areas, indicating a trend of increasing geographical inequality with respect to wage dynamics also from a within-region perspective. The range of ARLOP incidence values is wider where the median and average incidence is the highest, namely in Southern regions, with Sardinia and Sicily reporting a difference of over 100% between LLS areas with minimum and maximum incidence across all three definitions of AROLP. Interestingly, however, we observe an increase in inequality among LLS areas and also within northern regions, especially in the cases of Trentino Alto Adige, Emilia Romagna, Veneto and Liguria.

Overall our analysis shows how, over the course of a decade, the dualism which already characterised many aspects of the Italian economy has widened further with regard to the phenomenon of in-work poverty and low pay. Southern areas, where the incidence of low wages was already the highest, overwhelmingly reported the largest increases in AROLP rates. This trend emerges even clearer when we consider the incidence of low wages as the result of low work intensity and discontinuous employment, suggesting how the lack of stable employment offers presents a strong regional characterisation. In addition to an increasing wage inequality across the country's main regions, the analysis also highlighted a widening of within-region inequality, affecting Northern, Central and Southern regions alike.

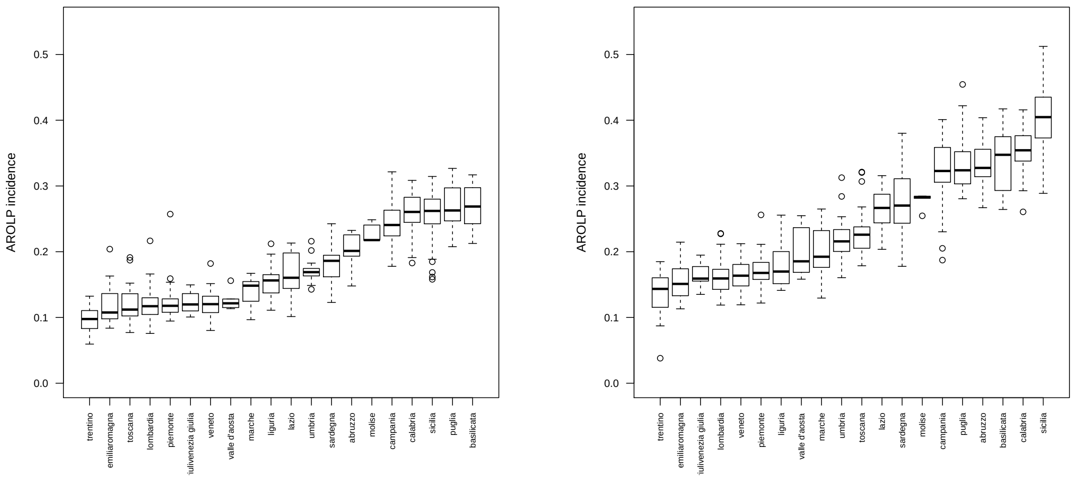


Figure 2.8: AROLP estimates based on hourly income at LLS level by region for the years 2008 (left panel) and 2017 (right panel)

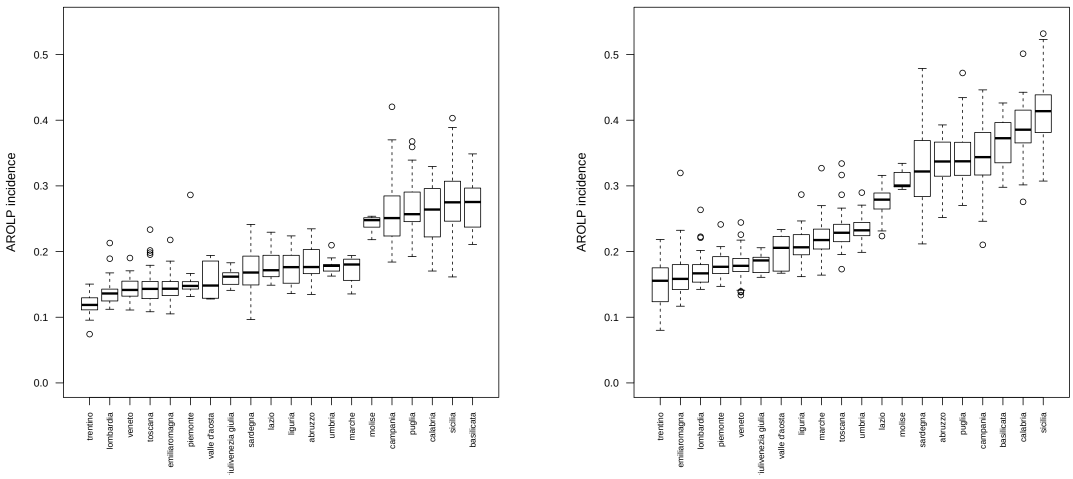


Figure 2.9: AROLP estimates based on weekly income at LLS level by region for the years 2008 (left panel) and 2017 (right panel)

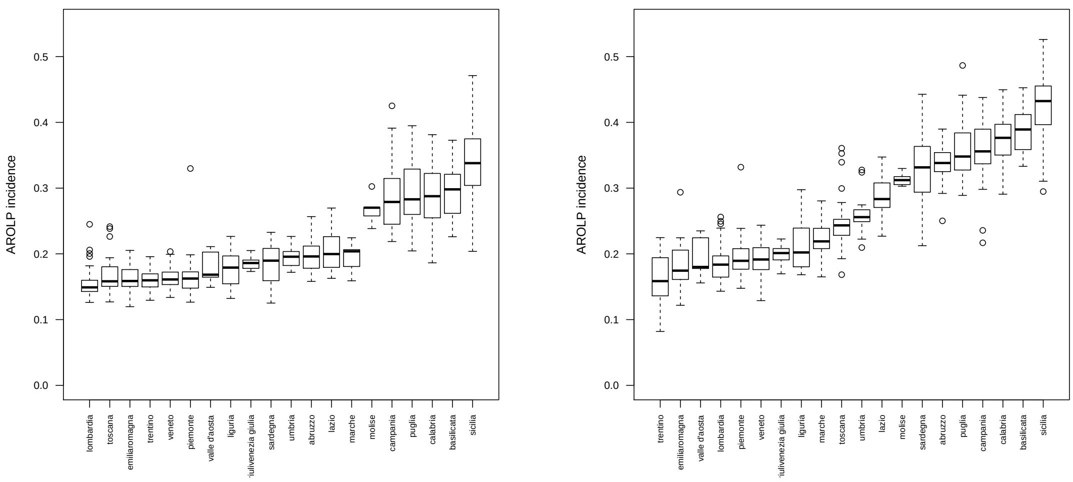


Figure 2.10: AROLP estimates based on annual income at LLS level by region for the years 2008 (left panel) and 2017 (right panel)

2.5 Sectoral employment correlation analysis

2.5.1 Panel fixed effect model

In order to explore the link between shifts in local sectoral employment and low-pay dynamics we construct a panel dataset at the LLS level. We consider the small area estimates for AROLP indicators for the years 2008, 2015, 2016 and 2017 as our dependent variables, obtained by applying the SAE method described in Subsection 2.4.1 to EU-SILC data. For each LLS we use the ASIA register data on sectoral employment as covariates. We thus exploit the longitudinal nature of our data to estimate the correlation between the employment variables and the three proposed ARLOP indicators by means of the following model specification for each definition of the ARLOP indicator (Annual, Weekly and Hourly)

$$Y_{d,m,t} = \beta \times SectEmpl_{m,t} + Empl_{m,t} + \gamma_m + \epsilon_m, \tag{2.6}$$

Where t=2008,2015, 2016, 2017, $Y_{m,t}$ is the AROLP rate for each of the three definitions d in area m at time t, $SectEmpl_{m,t}$ is an $m \times n$ matrix describing employment rate across 8 industrial sectors in each area m at time t. $Empl_{m,t}$ is an $m \times n$ matrix describing the overall level of employment, the proportion of payroll employees and self-employed in area m at time t, γ_p is a dummy variable equal to 1 when a region is equal to p. In this specification, the regional dummies act as time-invariant fixed effects and are instrumental in estimating the effect of intrinsic yet unobservable characteristics of individual regions. By including the region-level fixed effects we are somehow addressing the problem of omitted variable bias, even though the estimates of our model can not be interpreted in a causal way. Yet, this specification allows for exploring the potential link between them in a robust way.

2.5.2 Cross-section analysis

After having mapped the distribution of AROLP incidence across Italy's 610 LLS, we explore how this phenomenon relates to the local sectoral employment composition. Building on the results of the SAE models, we move to determine the cross-section correlation between the AROLP indicators and the employment-related variables contained in our dataset, and how this changed over the period 2008-2017. Table 2.6 reports the result of the OLS models for the years 2008 and 2017, where we regress the AROLP indicators for all three definitions of the employment-related variables. We observe how a positive and statistically significant correlation between the risk of low pay and the share of active population employed in agriculture and in non-market services persists throughout the period of observation. On the contrary, a persistent negative and statistically significant association is observed in relation to the local presence of jobs in the manufacturing sector, in transport and logistics and in areas where payroll employees and overall employment are greater. For other sectors, we observe how the association with the incidence of AROLP measure varied either across the years or across the different definitions considered in the analysis. For the construction sectors, for example, the association with AROLP incidence defined on annual income changed from statistically significant and negative in 2008 to statistically significant and positive in 2017. The share of active population employed in commerce showed no significant association with any of the AROLP definitions in 2008 to display a positive correlation with AROLP incidence based on hourly and annual income in 2017. A similar trend, but with an inverse dynamic, is observed in relation to the local share of professions. Interestingly, and somehow unexpectedly based on the literature on the topic, we observe how the hospitality sector has become negatively and statistically significantly associated with AROLP incidence across all three definitions at the end of the period of observation of this study.

Table 2.6: Cross-sectional analysis

VARIABLES	(2008) AROLP hourly	(2017) AROLP hourly	(2008) AROLP weekly	(2017) AROLP weekly	(2008) AROLP annual	(2017) AROLP annual
Agriculture	0.345***	0.297***	0.294***	0.302***	0.271***	0.267***
	(0.0460)	(0.0691)	(0.0536)	(0.0658)	(0.0587)	(0.0659)
Manufacturing	-0.0489***	-0.109***	-0.0360	-0.0909***	-0.0433*	-0.108***
	(0.0188)	(0.0323)	(0.0220)	(0.0308)	(0.0240)	(0.0308)
Constructions	-0.0367	0.238*	-0.192**	0.0742	-0.184**	0.242**
	(0.0637)	(0.123)	(0.0742)	(0.117)	(0.0812)	(0.117)
Commerce	-0.00324	0.210**	0.0252	0.0987	0.000821	0.226***
	(0.0600)	(0.0862)	(0.0699)	(0.0822)	(0.0766)	(0.0822)
Trans. and Log.	-0.366***	-0.341***	-0.389***	-0.257**	-0.510***	-0.335***
	(0.0928)	(0.127)	(0.108)	(0.121)	(0.118)	(0.121)
Hospitality	-0.0226	-0.112***	-0.00648	-0.110***	0.00848	-0.153***
	(0.0234)	(0.0360)	(0.0273)	(0.0343)	(0.0299)	(0.0343)
Professions	-0.119	-0.384	0.273	-0.742***	0.215	-0.777***
	(0.176)	(0.247)	(0.205)	(0.235)	(0.224)	(0.236)
Admin and Support serivces	-0.523***	-0.274*	-0.523***	-0.293**	-0.736***	-0.270*
	(0.118)	(0.146)	(0.138)	(0.139)	(0.151)	(0.139)
Health and Social Care	-0.470***	-0.384**	-0.528***	-0.198	-0.460**	-0.409**
	(0.152)	(0.169)	(0.178)	(0.161)	(0.194)	(0.161)
Non-market services	0.102***	0.225***	0.176***	0.221***	0.309***	0.225***
	(0.0361)	(0.0610)	(0.0421)	(0.0582)	(0.0460)	(0.0582)
Payroll	-0.213***	-0.316***	-0.314***	-0.274***	-0.262***	-0.373***
	(0.0302)	(0.0472)	(0.0352)	(0.0449)	(0.0385)	(0.0450)
Employment	-0.783***	-0.812***	-0.554***	-0.887***	-0.608***	-0.789***
	(0.0563)	(0.0521)	(0.0656)	(0.0496)	(0.0719)	(0.0497)
Constant	1.016***	1.064***	0.844***	1.152***	0.871***	1.116***
	(0.0575)	(0.0560)	(0.0671)	(0.0534)	(0.0734)	(0.0534)
Observations	610	610	610	610	610	610
R-squared	0.750	0.750	0.637	0.790	0.661	0.776

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Notes: Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. The dependent variable is the AROLP indicator based on hourly earnings (1 and 2), weekly earnings (3 and 4) and annual earnings (5 and 6). All independent variables are expressed as the share of the active population in each SLL.

2.5.3 Panel data analysis

Pooling these data across multiple time observations allows moving beyond the static cross-section correlation to estimate how the decline and growth of certain economic sectors at the local level are associated with the observed trends of low pay. Table 2.7 reports the findings of our panel data multivariate regression analysis. The findings portray a clear picture with regard to the trends related to variations in the incidence of AROLP and shift in sectoral employment. Over the years 2008 and 2017, the phenomenon of in-work poverty, defined as the proportion of workers at risk of low pay, has increased where the share of workers in the agriculture and commerce sectors has increased. Such correlation is observed regardless of whether the AROLP indicator is defined based on hourly, weekly or annual earnings. On the contrary, we observe a consistently negative and statistically significant correlation between our dependent variables and a range of sectors and employment characteristics. These include the share of manufacturing jobs, the share of jobs in hospitality, the share of jobs in professional and technical activities, and the share of jobs in administration and support services. With the exceptions of the transport and logistic sector displaying a negative and statistically significant coefficient only with the AROLP measure based on annual income, and with the health and social care sector displaying a statistically significant coefficient with the measure based on hourly and weekly income, our model shows overall consistency across the three definitions of AROLP. These findings are robust to a more expansive model specification, including per-capita Gross value Added data as a proxy of sectoral productivity, available at the province (NUTS 3 level). Table 2.C.1 in Appendix 2.C reports the findings of this robustness check.

Further, we observe the highest (negative) correlation coefficient in relation to the share of overall employment. In order to interpret these findings in light of the upward trend of in-work poverty recorded at the national level, it is important to consider the aggregate variations observed between 2008 and 2017 in our independent variables. The share of the overall labour force in employment, for example, decreased from 93.5% to 87%, coinciding with the doubling of the unemployment rate, meaning that such reduction is associated with a 1pp increase in AROLP incidence if annual earnings are considered. Our data shows that such a decrease follows an uneven pattern across Italy's main geographical areas, in line with the trend observed in the discussion of the AROLP findings. While northern areas, indeed, recorded on average a drop from 96.7% to 92.7% in the share of the overall workforce in employment (a 4.1% reduction), Central and Southern areas see this share drop from 95.4% to 89.9% (a 5.7% reduction) and from 90.1% to 81.0% (a 10% drop) respectively.

Similarly, the share of the total workforce employed in the manufacturing sector decreased significantly from 18.3% to 14.1%, with areas in the Centre and South regions recording the highest drop (23.4 and 25.6% respectively). These findings highlight two important aspects related to employment dynamics and wages. First, the strong negative association between overall employment and AROLP incidence confirms the findings of recent empirical work highlighting the stronger sensitivity of wages to unemployment at the bottom end of the wage distribution (Gregg and Machin, 2012; Gregg et al., 2014). Those who lose their jobs, indeed, are disproportionally represented by lower paid and lower educated workers (Faggio and Nickell, 2005). As such, in the context of high unemployment, the downward pressure on wages, resulting from increasing competition for fewer jobs, will affect predominantly lower-skilled and lower-paid workers. Thus, the economic downturn and slow recovery experienced by Italy during the period considered in this analysis not only resulted in the doubling rate of unemployment but is also reflected in record high rates of low-wage incidence. Second, these findings partially confirm the evidence that the decline of specific economic sectors is more strongly associated with the increase in low-wage incidence than others. The literature on the polarization of labour markets points to the decline of the manufacturing sectors and to the rise of low-skilled service jobs as the source of increase in low wage incidence (David and Dorn, 2013; Cormier and Craypo, 2000). The findings of our analysis in the context of Italy, where manufacturing jobs constituted both the largest share of the workforce and the sector that recorded the largest decline in the period considered, provide further evidence for this argument.

Table 2.7: Panel-data fixed effects model

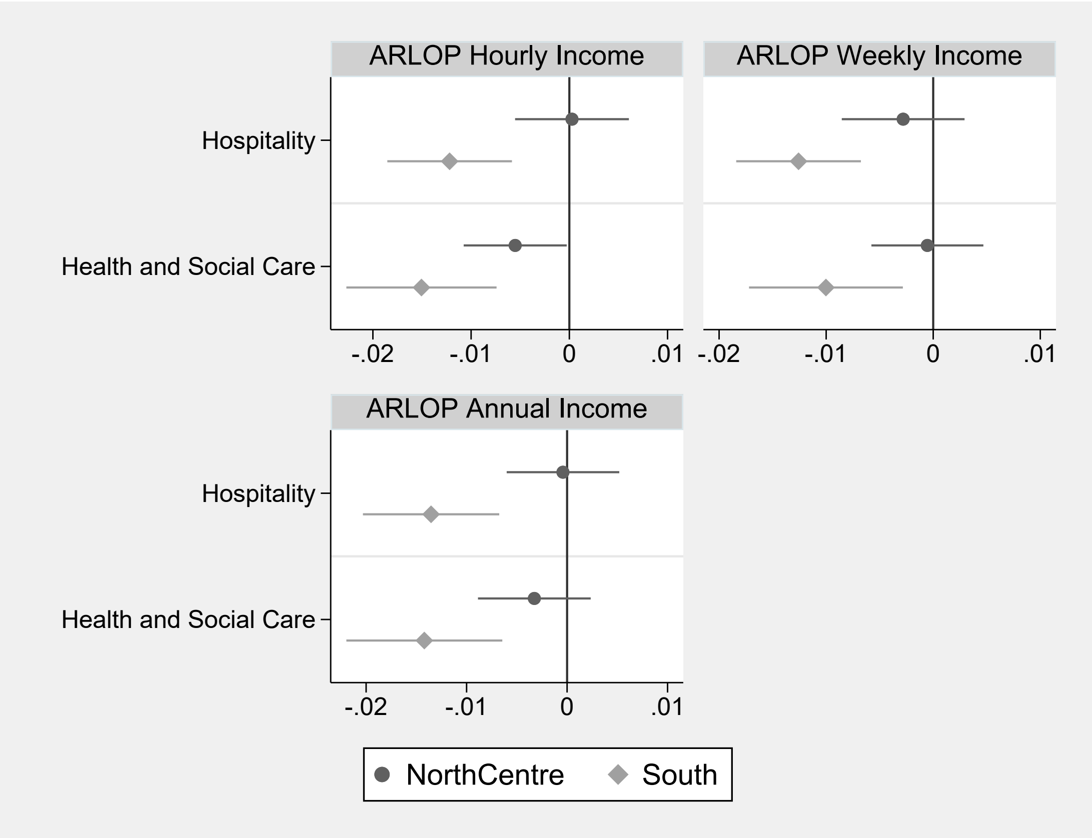
VARIABLES	(1) AROLP hourly income	(2) AROLP weekly income	(3) AROLP annual income
Agriculture	0.00882***	0.00505***	0.00712***
	(0.00221)	(0.00195)	(0.00202)
Manufacturing	-0.0106***	-0.0106***	-0.0121***
	(0.00224)	(0.00213)	(0.00225)
Constructions	-0.000454	-0.00363	-0.00268
	(0.00355)	(0.00337)	(0.00374)
Commerce	0.0257***	0.0257***	0.0258***
	(0.00647)	(0.00540)	(0.00664)
Trans. and Log.	-0.00350*	-0.00368*	-0.00597**
	(0.00210)	(0.00215)	(0.00234)
Hospitality	-0.00454**	-0.00665***	-0.00642***
	(0.00221)	(0.00205)	(0.00224)
Professions	-0.0147***	-0.0159***	-0.0170***
	(0.00432)	(0.00430)	(0.00461)
Admin and Support serivces	-0.00768***	-0.00796***	-0.00948***
	(0.00160)	(0.00168)	(0.00172)
Health and Social Care	-0.00843***	-0.00306	-0.00626**
	(0.00241)	(0.00236)	(0.00255)
Non-market services	-0.0261***	-0.0209***	-0.0132
	(0.00833)	(0.00800)	(0.00834)
Employment	-0.129***	-0.188***	-0.173***
	(0.0303)	(0.0314)	(0.0343)
Payroll	-0.272***	-0.262***	-0.245***
	(0.0264)	(0.0270)	(0.0286)
Constant	0.126***	0.129***	0.127***
	(0.0264)	(0.0250)	(0.0264)
Observations	2,440	2,440	2,440
Number of sll_2011	610	610	610
Region fixed-effects	Yes	Yes	Yes
Year fixed-effects	Yes	Yes	Yes

Notes: Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. The dependent variable is the AROLP indicator based on hourly earnings (1), weekly earnings (2) and annual earnings (3). All independent variables are expressed as the share of the active population in each SLL and are log-transformed to obtain a better linear approximation.

When it comes to services, however, the picture emerging from the analysis is rather less clear. While the growth of the commerce sector shows a positive association with the incidence of AROLP across all three definitions considered, the findings for other service sectors, such as health and social care and hospitality appear at odds with what the literature suggests, presenting a negative correlation between their growth and that of low-pay incidence. These sectors indeed are characterised by low union representation, low productivity and use of short-term employment contracts, all features typically associated with lower salaries and the presence of in-work poverty (Dwyer, 2013; Vanselow et al., 2010). Additional regression results obtained by applying the same fixed effect panel model to LLS areas located in the South and in the Northern and Central regions separately provide interesting insights to unpack the dynamics described here. Figure 2.11 displays how the negative coefficients in the share of active population employed in hospitality and health and social care observed at the national level are driven largely by southern areas, while no statically significant association is found in the other areas of the country. Interestingly, both sectors represent the only ones to record a slight increase in the share of active population employed within southern regions, where the overall fall in employment has been the sharpest.

Based on the current data available it is difficult to establish whether we are witnessing a crowding-in effect into service sectors as the combined result of processes of de-industrialisation and contraction of labour demand, or an over-representation of highly paid jobs in these sectors, notoriously characterised by informal employment arrangements not captured in the data. Access to worker's microdata has the potential to shed new light on wage dynamics related to processes of de-industrialisation coupled with economic contraction in depressed economic areas such as Italy's southern regions.

Figure 2.11: Regression coefficients of Hospitality and Health and Social Care by macro areas



2.6 Conclusion

The study demonstrated the relevance and importance of the sub-regional analysis of in-work poverty. From an overall significant increase in the AROLP incidence at the national level between the years 2008 and 2017, the application of SAE methods unveils a strong geographical heterogeneity in the trends. Moreover, focusing on small areas such as LLS as the main unit of analysis allows for uncovering important trends across administrative boundaries as well as within-region heterogeneity. The findings highlight how a decade characterised by a deep economic recession and slow recovery has contributed to widening the country's north-south dualism in relation to the phenomenon of low pay. Southern areas, where the incidence of low wages was already the highest, overwhelmingly reported the largest increases in AROLP rates. This trend emerges even clearer when we consider the incidence of low wages as the result of low work intensity and discontinuous employment, suggesting how the lack of stable employment presents a strong regional characterisation. In addition to an increasing wage inequality across the country's main regions, the granularity of the analysis also highlighted a widening of within-region inequality, affecting Northern, Central and Southern regions alike.

The findings of the small area level geographical analysis allow us to associate the geographical heterogeneity in trends of ARLOP incidence with local sectoral employment characteristics. The empirical findings of our panel regression analysis reveal how areas characterised by growth in low-skill and low-productivity sectors such as agriculture and commerce are those that recorded higher increases in AROLP incidence. On the contrary trends of low pay are negatively associated with the growth of manufacturing jobs, admin and support services to enterprises, and technical and scientific professions. In addition, variations in overall employment represent the strongest predictor for dynamics of low-pay incidence. Interestingly, the panel correlation analysis unveiled trends at odds with the literature on sectoral employment and wages in relation to the sectors of hospitality and health and social care. Driven by areas in southern regions, the growth of these sectors in Italy is associated with a lower level of AROLP incidence. Access to geo-localised microdata on earnings and employment would allow unpacking the mechanisms behind these findings, casting light on the role that sectors traditionally associated with low productivity and low wages play against a backdrop of de-industrialisation and economic contraction as present in the Southern Italian context.

The findings of this research constitute novel and important evidence of the local distribution of low-pay and in-work poverty in Italy. National as well as local policymakers can rely on this information to better comprehend the phenomenon and target interventions. Moreover, while not establishing a clear causal link, the result of the correlation analysis lay out the effect on wages associated with the local decline and growth of key specific sectors, presenting useful evidence for the formulation of employment and industrial policies.

Appendix

2.A Direct estimates at region-level

Table 2.A.1: 2008 direct estimates at region-level

Region	Domain	AROLP hourly	CV	Direct weekly	CV	AROLP annual	CV
Piemonte	ITC1	11.40%	9.75%	14.74%	8.99%	15.69%	8.66%
Valle d'Aosta	ITC2	11.12%	18.91%	13.22%	17.45%	15.19%	16.10%
Lombardia	ITC3	14.97%	11.65%	16.52%	11.02%	16.83%	10.88%
Trentino Alto Adige	ITC4	10.87%	7.44%	13.65%	6.67%	14.89%	6.26%
Veneto	ITD3	11.18%	7.86%	14.21%	6.99%	15.96%	6.70%
Friuli Venezia-Giulia	ITD4	12.80%	11.50%	16.70%	9.87%	18.74%	9.36%
Liguria	ITD5	10.56%	8.86%	14.21%	7.61%	15.82%	7.17%
Emilia Romagna	ITDA	8.29%	12.09%	12.14%	10.48%	15.35%	9.19%
Toscana	ITE1	10.47%	10.32%	13.84%	8.66%	15.39%	8.05%
Umbria	ITE2	16.29%	10.18%	17.40%	9.48%	18.50%	9.18%
Marche	ITE3	13.03%	9.17%	16.79%	8.14%	18.89%	7.78%
Lazio	ITE4	12.03%	8.62%	15.59%	7.62%	17.24%	7.18%
Abruzzo	ITF1	19.48%	13.47%	17.51%	13.76%	18.91%	13.27%
Molise	ITF2	22.66%	13.38%	23.58%	13.22%	25.72%	12.59%
Campania	ITF3	22.38%	8.38%	21.83%	8.42%	23.95%	8.15%
Puglia	ITF4	24.78%	7.98%	24.17%	7.86%	26.00%	7.55%
Basilicata	ITF5	25.37%	12.23%	24.76%	12.03%	26.68%	11.54%
Calabria	ITF6	24.23%	11.12%	22.81%	10.69%	25.74%	10.20%
Sicilia	ITG1	22.15%	9.40%	23.49%	9.03%	28.52%	8.01%
Sardegna	ITG2	15.61%	14.12%	13.72%	14.50%	15.20%	13.71%

Table 2.A.2: 2017 direct estimates at region-level

Region	Domain	AROLPhourly	CV	AROLPweekly	CV	AROLPannual	CV
Piemonte	ITC1	16.27%	8.08%	17.11%	8.11%	18.01%	7.88%
Valle d'Aosta	ITC2	16.82%	14.38%	17.56%	14.59%	18.06%	14.64%
Lombardia	ITC3	13.69%	6.60%	15.83%	6.23%	16.96%	5.99%
Trentino Alto Adige	ITC4	13.13%	11.04%	14.36%	10.34%	15.87%	10.17%
Veneto	ITD3	16.78%	6.80%	17.77%	6.60%	18.98%	6.41%
Friuli Venezia-Giulia	ITD4	15.90%	8.91%	18.03%	8.39%	19.05%	8.22%
Liguria	ITD5	16.88%	9.51%	20.98%	8.76%	20.88%	8.74%
Emilia Romagna	ITDA	14.39%	7.74%	15.58%	7.34%	16.63%	7.12%
Toscana	ITE1	21.70%	6.84%	22.54%	6.70%	22.87%	6.63%
Umbria	ITE2	20.13%	10.52%	22.65%	9.49%	23.68%	9.24%
Marche	ITE3	18.09%	9.00%	20.81%	8.26%	21.18%	8.14%
Lazio	ITE4	22.44%	6.36%	24.08%	6.09%	24.65%	5.98%
Abruzzo	ITF1	31.55%	10.90%	32.19%	10.74%	31.91%	10.70%
Molise	ITF2	26.00%	12.65%	28.46%	12.03%	29.15%	11.89%
Campania	ITF3	29.80%	7.87%	31.91%	7.60%	32.35%	7.60%
Puglia	ITF4	30.98%	7.53%	32.26%	7.32%	32.93%	7.25%
Basilicata	ITF5	30.15%	11.59%	33.51%	10.91%	35.93%	10.65%
Calabria	ITF6	33.89%	9.41%	35.94%	9.35%	35.73%	9.29%
Sicilia	ITG1	36.73%	7.40%	37.79%	7.35%	38.01%	7.28%
Sardegna	ITG2	21.82%	11.34%	27.15%	10.46%	27.47%	10.27%

2.B SAE estimates result for EU-SILC waves of 2015 and 2016

Table 2.B.1: Coefficients of variation, Brown test p values and correlation of at-risk-of low pay estimates (annual earnings), obtained with FH arcsine models on EUSILC 2015 AND 2016. The estimates were benchmarked at the regional level using the direct estimates for the region as benchmark.

2015	<16.5%	16.5-33.3%	>33.3%	Brown test	Corr. Direct
$AROLP_{annual}$	600	10	0	p=0.252	0.56
$AROLP_{weekly}$	603	7	0	p=0.293	0.55
$AROLP_{hourly}$	334	273	4	p=0.999	0.62
2016					
$AROLP_{annual}$	532	75	3	p=0.597	0.61
$AROLP_{weekly}$	544	63	3	p=0.707	0.685
$AROLP_{hourly}$	339	267	4	p=0.373	0.60

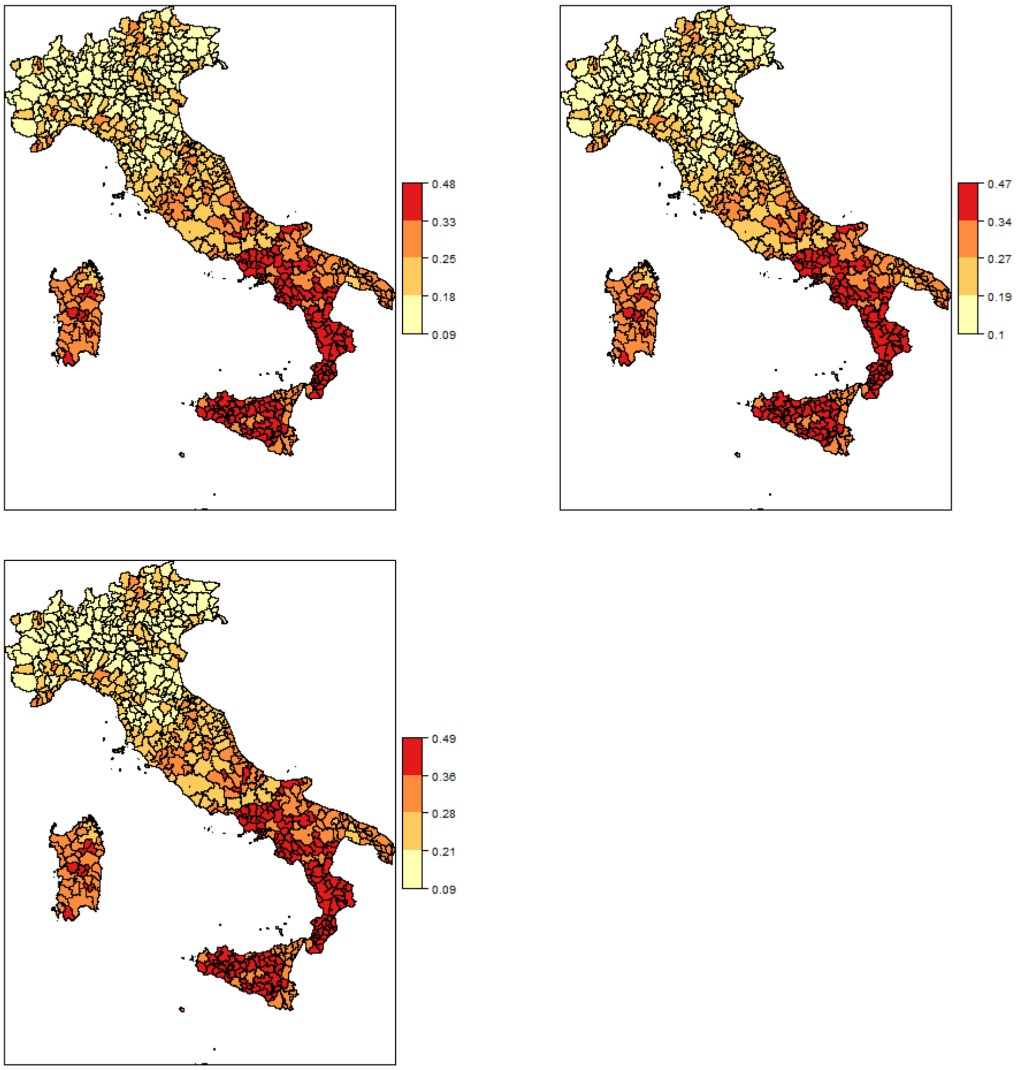


Figure 2.B.1: At-risk-of low pay rates based on EU-SILC 2015 for annual income (panel a), weekly income (panel b) and hourly income (panel c)

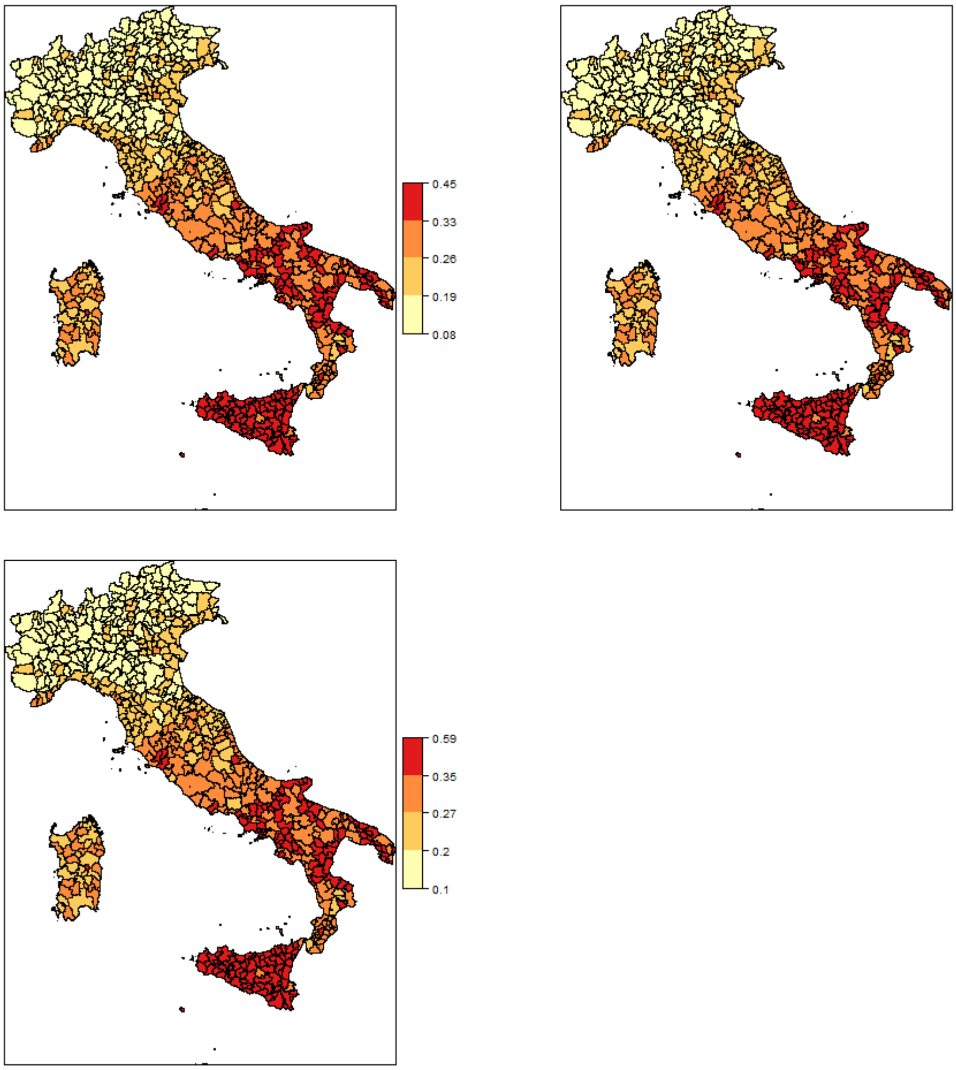


Figure 2.B.2: At-risk-of low pay rates based on EU-SILC 2016 for annual income (panel a), weekly income (panel b) and hourly income (panel c)

2.C Additional regression tables

Table 2.C.1: Panel regression with GVA pc controls

VARIABLES	(1) fh_hourly_benched	(2) fh_weekly_benched	(3) fh_annual_benched
Agriculture	0.00831***	0.00523**	0.00677***
	(0.00238)	(0.00233)	(0.00242)
Manufacturing	-0.0139***	-0.0147***	-0.0174***
	(0.00219)	(0.00221)	(0.00228)
Constructions	0.00213	0.000567	0.00177
	(0.00367)	(0.00362)	(0.00408)
Commerce	0.0237***	0.0205***	0.0168***
	(0.00649)	(0.00566)	(0.00647)
Trans. and Log.	-0.00323	-0.00412*	-0.00640***
	(0.00224)	(0.00237)	(0.00248)
Hospitality	-0.00446*	-0.00662***	-0.00727***
	(0.00235)	(0.00230)	(0.00246)
Professions	-0.0166***	-0.0178***	-0.0187***
	(0.00439)	(0.00445)	(0.00467)
Admin. and Support services	-0.0104***	-0.0110***	-0.0122***
	(0.00165)	(0.00176)	(0.00179)
Health and Social Care	-0.00639***	-0.000853	-0.00467*
	(0.00248)	(0.00250)	(0.00276)
Non-market services	-0.0116	-0.0106	-0.00481
	(0.00933)	(0.00977)	(0.0103)
Employment	-0.0661**	-0.137***	-0.123***
	(0.0336)	(0.0356)	(0.0363)
Payroll	-0.0853***	-0.0550***	-0.00598
	(0.0125)	(0.0134)	(0.0142)
GVA pc Agriculture	-0.000129*	-5.55e-05	-7.65e-05
	(6.68e-05)	(7.02e-05)	(7.91e-05)
GVA pc Industry	-0.000385***	-0.000347***	-0.000384***
	(0.000101)	(0.000103)	(0.000102)
GVA pc Construction	6.20e-05	0.000209	0.000561**
	(0.000212)	(0.000241)	(0.000248)
GVA pc Retail, Transport, Hospitality	-0.000636***	-0.000224	-0.000393**
	(0.000184)	(0.000180)	(0.000190)
GVA pc Financial and Business Services	-0.000238**	-7.96e-05	-6.52e-05
	(9.99e-05)	(9.59e-05)	(0.000100)
GVA pc Non-market services	-0.000414	-0.000966**	-0.00177***
	(0.000414)	(0.000394)	(0.000431)
Constant	0.0951***	0.0696**	0.0938***
	(0.0327)	(0.0329)	(0.0330)
Observations	2,440	2,440	2,440
Number of sll_2011	610	610	610
Region fixed-effects	Yes	Yes	Yes
Year fixed-effects	Yes	Yes	Yes

Robust standard errors in parentheses
*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Notes: Standard errors in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$. The dependent variable is the AROLP indicator based on annual earnings (1), weekly earnings (2) and hourly earnings (3). All independent variables are expressed as the share of the active population in each SLL and are log-transformed to obtain a better linear approximation.

Table 2.C.2: Panel regression with SLL fixed effect

VARIABLES	(1) AROLP hourly income	(2) AROLP weekly income	(3) AROLP annual income
Agriculture	0.0120	-0.00776	-0.000607
	(0.00820)	(0.0103)	(0.00925)
Manufacturing	-0.00172	-0.0179*	-0.0255***
	(0.00741)	(0.00965)	(0.00831)
Constructions	-0.0263***	-0.0389***	-0.0390***
	(0.00855)	(0.0105)	(0.00869)
Commerce	0.00767	0.0100	0.00423
	(0.0161)	(0.0186)	(0.0182)
Trans. and Log.	0.00751*	0.00809	0.00674
	(0.00437)	(0.00597)	(0.00551)
Hospitality	0.00499	0.000924	0.00118
	(0.00727)	(0.0102)	(0.00851)
Professions	-0.000288	-0.0149	-0.0126
	(0.00942)	(0.0133)	(0.0114)
Admin. and Support services	-0.00513**	-0.00533	-0.00439
	(0.00258)	(0.00359)	(0.00322)
Health and Social Care	-0.0125***	-0.00374	-0.00450
	(0.00459)	(0.00656)	(0.00564)
Non-market services	-0.0380**	-0.0162	0.00694
	(0.0159)	(0.0173)	(0.0145)
Employment	-0.187***	-0.366***	-0.357***
	(0.0412)	(0.0518)	(0.0442)
Payroll	-0.127***	-0.0456	0.0187
	(0.0275)	(0.0297)	(0.0260)
Constant	0.0541	-0.0492	0.0141
	(0.0547)	(0.0739)	(0.0666)
Observations	2,440	2,440	2,440
R-squared	0.603	0.530	0.518
Number of sll_2011	610	610	610
LLS fixed-effects	Yes	Yes	Yes
Year fixed-effects	Yes	Yes	Yes

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 2.C.3: Panel regression by macro areas

VARIABLES	(1) North&Centre	(2) South	(3) North&Centre	(4) South	(5) North&Centre	(6) South
Agriculture	0.0139***	0.00815**	0.00839***	0.00806**	0.0115***	0.00891**
	(0.00241)	(0.00391)	(0.00209)	(0.00349)	(0.00214)	(0.00401)
Manufacturing	-0.0154***	-0.00123	-0.0165***	-0.000625	-0.0147***	-0.00437
	(0.00318)	(0.00392)	(0.00349)	(0.00342)	(0.00320)	(0.00394)
Constructions	-0.00149	0.00737	-0.00473	0.00610	-0.00122	0.00869*
	(0.00433)	(0.00478)	(0.00473)	(0.00438)	(0.00466)	(0.00511)
Commerce	0.00309	0.0280***	0.00902	0.0176**	0.00280	0.0211**
	(0.00690)	(0.00875)	(0.00713)	(0.00710)	(0.00756)	(0.00858)
Trans. and Log.	-0.00141	-0.0136***	-0.00117	-0.0139***	-0.00168	-0.0182***
	(0.00285)	(0.00360)	(0.00312)	(0.00340)	(0.00312)	(0.00372)
Hospitality	0.00426	-0.0127***	0.000454	-0.0139***	0.00260	-0.0160***
	(0.00330)	(0.00335)	(0.00336)	(0.00313)	(0.00323)	(0.00348)
Professions	-0.00744*	-0.0243***	-0.00819*	-0.0293***	-0.00629	-0.0329***
	(0.00440)	(0.00674)	(0.00461)	(0.00639)	(0.00515)	(0.00641)
Admin. and Support services	-0.00544***	-0.00939***	-0.00599***	-0.00874***	-0.00589***	-0.0109***
	(0.00195)	(0.00212)	(0.00210)	(0.00224)	(0.00209)	(0.00214)
Health and Social Care	-0.00321	-0.0141***	0.00135	-0.00918**	-0.00138	-0.0142***
	(0.00251)	(0.00411)	(0.00255)	(0.00393)	(0.00276)	(0.00415)
Non-market services	-0.0125	-0.0191	-0.0219***	0.0175	-0.0171**	0.0177
	(0.0103)	(0.0138)	(0.00788)	(0.0147)	(0.00837)	(0.0142)
Employment	-0.271***	0.0888**	-0.241***	0.00841	-0.261***	0.0539
	(0.0674)	(0.0384)	(0.0650)	(0.0352)	(0.0725)	(0.0348)
Payroll	-0.130***	-0.0560***	-0.108***	-0.0226	-0.0989***	0.0292*
	(0.0197)	(0.0155)	(0.0222)	(0.0148)	(0.0214)	(0.0167)
Constant	0.0414		0.0394		0.0740***	
	(0.0294)		(0.0286)		(0.0269)	
Observations	1,320	1,120	1,320	1,120	1,320	1,120
Number of sll_2011	330	280	330	280	330	280
Region fixed-effects	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed-effects	Yes	Yes	Yes	Yes	Yes	Yes

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Chapter 3

Disaggregation of poverty indicators by small area methods for assessing the targeting of the "Reddito di Cittadinanza"

Abstract: In Italy, a crucial anti-poverty policy "Reddito di Cittadinanza" (RdC), a measure of guaranteed minimum income, was introduced in April 2019. We aim to evaluate the targeting of the RdC policy at the local level, as aggregated analyses could mask important misalignments between the share of beneficiaries of the RdC and the share of poor households. To measure the poverty share in the local areas of interest, two main indicators to capture and monitor poverty are used in Europe: the At-Risk-of-Poverty Rate based on the EU Statistics on Income and Living Conditions survey and the Absolute Poverty Index based on consumption data collected through the Household and Budget Survey. To obtain reliable estimates of these indicators at the local level, it is necessary to introduce small area estimation models that allow the use of data from different sources. We apply a bivariate Fay and Herriot model to provide reliable estimates of absolute and relative poverty for the assessment of RdC policy targeting in the 59 areas represented by the region by degree of urbanisation level in Italy. The degree of urbanisation is indeed a key geographical variable in the study of the poverty phenomenon. Our results suggest that the RdC policy implemented at the national level shows heterogeneous targeting performance at the local level, excluding large shares of poor households from the program. These findings yield a set of policy implications for improving the targeting of the measure.

3.1 Introduction

In April 2019, the Italian government introduced a national measure of guaranteed minimum income under the name of "Reddito di Cittadinanza" (RdC) or Citizens' Income. With regard to both the cohort of beneficiaries and the level of monetary support available, the RdC represents the largest monetary transfer program for low-income families in the history of the Italian social security system (Baldini and Gori, 2019), which, as recently as 2017, did not present any form of guaranteed minimum income scheme. For the year 2019 alone, the total program expenditure was forecasted at €5.6bn, with an estimated cohort of beneficiaries of 1.3m households. Amidst the background of economic stagnation and household income contraction, the measure was introduced with the objectives of (i) tackling poverty and reducing inequality, (ii) increasing labour market participation, and (iii) improving social inclusion. Official statistics by the Italian National Statistical Institute (ISTAT) show how the number of households in absolute poverty in Italy has been on the rise over the five years prior to the introduction of the policy. In 2018, the number of families in absolute poverty reached 1.8m, with an absolute poverty incidence of 7% (ISTAT, 2019). Against this backdrop, the government's estimate indicated a total of 1.3m households as potential beneficiaries of the RdC (DDL n.1018, 2019). This figure was later adjusted to 960,000 beneficiary households as the policy reached its fifth month of implementation, highlighting a gap between the overall cohort of RdC beneficiaries and the total number of families in absolute poverty in Italy (compare INPS, 2019). Preliminary evaluation studies conducted estimated the total number of beneficiary households to be between 1.1m and 1.4m (Monducci, 2019; Boeri, 2019a; Baldini et al., 2002; Curci et al., 2020), significantly below the latest available estimates for absolute poverty of 1.8m. Scholars have identified specific aspects in the design of the policy with the potential to hinder the effectiveness of the scheme in reducing poverty. In particular, analyses have pointed to the novel equivalence scale adopted by the regulation and the 10-year residence eligibility criterion as features which might limit the targeting performance of the RdC (Baldini and Gori, 2019; Boeri, 2019a; Curci et al., 2020; Fierro, 2019).

Based on these considerations, on the government forecasts and official statistics, several questions arise regarding the stated objectives of the policy. To what extent does the RdC succeed in targeting support to families in poverty? How does the cohort of identified beneficiaries of the RdC map against the distribution of families living in poverty in Italy at the national, regional, and sub-regional levels? What are the policy's redistributive effects on the incidence of poverty in the country? Which factors related to local demographic and economic characteristics drive variations in targeted coverage and take-up rates among different geographical areas?

Capturing geographical heterogeneity in the effects of anti-poverty interventions is a key requirement for policymakers and researchers. In a country like Italy, characterised by strong economic dualism between its northern and southern regions, the geographical distribution of poverty follows clear regional and sub-regional patterns (ISTAT, 2019), as discussed in Section 1.1. As such, a comprehensive assessment of RdC targeting performance should consider its heterogeneous performance across regional divides. In addition, a vast body of literature considers the degree of urbanisation as a key geographical variable in the study of poverty phenomena (Satterthwaite and Tacoli, 2002; Weziak-Bialowolska, 2016). Levels of poverty in rural and urban areas not only differ (Visaria, 1980; Tanton et al., 2010) but they stem from different causes and require different solutions (Wang et al., 2012; Atkinson et al., 2010). Therefore, in this study, we consider the degree of urbanisation, as captured in the DEGURBA classification by Eurostat (Union et al., 2021), as the geographical unit of our analysis across each of Italy's 20 regions.

When relying on data from national sample surveys, however, expanding the level of geographical detail of the estimates to the sub-regional level can be difficult, as at this level the variances of the estimates are often too high due to the design of the survey. While administrative tax and revenues registers represent the ideal data source for this type of analysis, their use by government agencies and researchers remains limited due to technical and legal barriers still in place (Pratesi and Salvati, 2016). As discussed in Section 1.3, by and large, official poverty indicators are estimated on the basis of surveys collected by national statistical agencies at the national level, and often, due to their limited sample sizes, cannot provide accurate estimates at lower sub-regional units of analysis (Tzavidis et al., 2018). Small Area Estimation (SAE) methods offer the tools to overcome this gap.

The first contribution of this study is to provide a new application of SAE methods to assess the targeting performance of an anti-poverty program. While SAE methods are commonly used to study the geographic distribution of poverty, in this application, this methodology is instrumental in providing baseline poverty estimates for each of the 59 areas of analysis to successively estimate the targeting performance of the RdC across such areas. Second, the research will provide the first assessment of RdC targeting based on administrative data. All existing studies focusing on this topic have been conducted using micro-simulation approaches based on survey data, thus relying on a set of assumptions about take-up and income reporting information (Baldini et al., 2002; Curci et al., 2020; Monducci, 2019). This has made them valuable for forecasting the impact of the policy ex-ante, but of limited use in assessing the scheme ex-post. By contrast, our study provides the first assessment of the actual targeting performance of the measure. The results and conclusions drawn by our research, therefore, provide useful evidence for policymakers to improve the design of crucial livelihood policies and to ensure the effective targeting of public funds toward contrasting poverty across all territories. The remainder of this chapter is structured as follows. Section 3.2 presents key provisions of the RdC scheme. Section 3.3 describes the data employed and discusses the definitions of poverty considered in the analysis. Section 3.4 presents the SAE models applied to the analysis, Section 3.5 discusses the main findings and Section 3.6 concludes the paper.

3.2 Institutional background

In this section, we provide the key elements of the RdC policy and present some of the existing literature analysing its design and provisions.

3.2.1 RdC key provisions

The main RdC provision is the introduction of a monthly cash handout for low-income households, whose eligibility requirements consist of a set of demographic and income-based criteria. There are, first of all, constraints of demographic nature: the claimant must have Italian or EU citizenship or hold an EU residence permit and reside in Italy for at least 10 years. With regard to income and capital requirements, the assessment is based on the score of the ISEE, an indicator of household economic circumstances equivalent to household size and specific needs (e.g., the presence of disability), and on the possession of high-value goods such as newly acquired motor vehicles and second homes. The RdC monetary benefit consists of two elements: first, a monetary support to household income, up to the threshold of €6,000 per year multiplied by the corresponding coefficient of equivalence, and second, a rent subsidy, up to a maximum of €3,360 per year (or €280 per month). The maximum amount of RdC for a single-member household in a rented property with no income is €9,360 per year (€780 per month). The amount of the benefit cannot be less than €480 per year (€40 per month) or not more than €15,960 per year (€1,330 per month for a family living in a rented property with no income and two or more adult dependants or four or more children). The legislation establishes that all adult members of a beneficiary household, with the exception of identified exemptions, must provide immediate availability to work and are called upon to sign an Employment Pact which consists of a personalised service pact for job placement or training. The signatories of the Employment Pact are required to accept at least one of three suitable job offers, where adequacy is defined with reference to previous legislation. For those families whose needs are identified as "complex" (DLn.4/2019), the measure requires the stipulation of a Social Inclusion Pact, a personalised program of social and welfare support, training, and employment activation.

3.2.2 Main concerns

Authors identified two main issues in the design of the policy with the potential to mitigate the effectiveness of the scheme in reducing poverty. The first issue refers to the choice of the equivalence scale employed to adjust the amount of benefit to the different sizes and needs of the households. The weights to be applied to any additional adult or child in the household, as established by the regulation, are indeed lower than the weights established by the so-called modified OECD scale, an international measure used by Eurostat, and in place in other income support schemes previously adopted in Italy (Monducci, 2019). As a result, the scheme overwhelmingly benefits households composed of single adults over larger families, especially when compared to the differences in benefit amounts under a counterfactual scenario employing the OECD-modified scale (Baldini et al., 2019; Curci et al., 2020). This contrasts with the demographic trends emerging from the official statistics on poverty published by ISTAT. The report highlights how large families (i.e., couples with three or more children) are more than twice as likely to live in absolute poverty compared to single adults, with a poverty incidence of 16.4% in the former group compared to 6.4% in the latter group (ISTAT, 2019). Building on this evidence, critics of the policy point at how the choice of the "peculiar equivalence scale" (Boeri, 2019a) adopted in the RdC might lead to a mismatch between the cohort of beneficiaries of the scheme and that of families living in poverty in Italy.

Second, concerns have been raised over the 10-year residence eligibility criteria in place for non-EU citizens. For pressure groups and commentators alike, this threshold is considered too strict (Baldini and Gori, 2019; Fierro, 2019; Curci et al., 2020). This measure excludes some of the most vulnerable demographic groups from a vital source of economic and social support. According to the national statistics on poverty published by ISTAT, the incidence of poverty among foreign residents is close to five times higher than that among Italian families (ISTAT, 2019). Essentially, current eligibility design leaves some of the demographic groups at a higher risk of poverty and social exclusion from the support framework of the measure, thus undermining the scheme in achieving its objectives (Boeri, 2019b).

3.3 Data and definitions

3.3.1 Data

The analyses presented in this study are based on four main data sources. First, estimates for absolute poverty (AP) were based on the Household Budget Survey (HBS) data for 2017. The survey was collected yearly by ISTAT and provided information on household consumption behaviour. The dataset provides a flag for households living in absolute poverty, and comprises approximately 17,000 observations. The estimates produced by the HBS were reliable at the regional level. Instead, estimates at the sub-regional level are characterised by high variability, as the sample size can be too small to obtain reliable results. Table 1 reports some statistics of the HBS sample size for the 59 areas identified by considering the Italian regions and DEGURBA classification. As we can see, the minimum sample size is equal to 40, while the first 25% of the areas have a sample size lower or equal to 134 households. Estimates for the at-risk-of-poverty rate (AROP) across all 59 areas of analysis were based on European Union Statistics on Income and Living Conditions (EU-SILC) survey data collected in 2017. The EU-SILC aims to collect timely and comparable cross-sectional and longitudinal multidimensional microdata on income, poverty, social exclusion, and living conditions in Europe. The 2017 wave of the survey contained information on self-reported income for 2016, with 22,200 observations. As for HBS, Table 3.3.1 reports some statistics of the EU-SILC sample size for the 59 areas of interest. As we can see, in this case, the minimum number of sampled households is equal to only 16 households.

Table 3.3.1: Descriptive statistics of the sample size of HBS and EU-SILC surveys across the 59 regions by DEGURBA areas.

	Min	1st Qu.	Median	3rd Qu.	Max
HBS	40.0	134.0	218.0	377.0	1066.0
EU-SILC	16.0	211.0	337.0	477.5	1254.0

Statistics Canada (Wannell and Usalcas, 2012) provides guidelines for publication related to the uncertainty of estimates; estimates with a coefficient of variation (CV) less than 16.6% are considered reliable for general use, estimates with CVs between 16.6% and 33.3% should be accompanied by warnings to users, and estimates with coefficients of variation greater than

33.3% are considered unreliable. Table 3.3.2 shows the number of areas classified according to these three thresholds of the CV on the basis of the 'direct' estimates - that is, estimates computed using only survey data - of the AROP and AP. As we can see, out of the 59 total areas of interest, the number of areas with a CV in the second or third class of CV values is rather high for both indicators, suggesting the need to resort to appropriate modelling techniques to reduce the estimated CVs.

Table 3.3.2: Coefficients of variation of AP and AROP direct estimates.

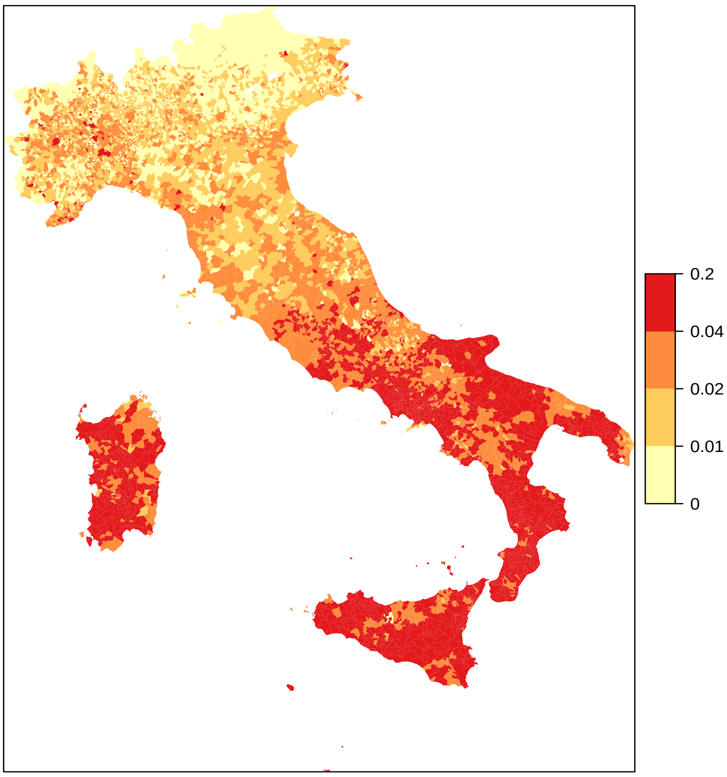
		<16.5%	16.5-33.3%	>33.3%
AP	Direct	5	32	22
AROP	Direct	33	24	2

The basic idea of the SAE methods is to introduce a statistical model to exploit the relationship between the variable of interest and some covariates for which population information is available to improve the precision of direct estimates. We consider as auxiliary variables for both the HBS and EU-SILC data, a set of administrative covariates from the Italian Ministry of Treasure Tax returns data referred to the year 2017 at the municipality level. The variables employed in the models are (i) the percentage of taxpayers, (ii) the percentage of payroll employees, (iii) the percentage of the population with yearly income below €10,000, (iv) the percentage of the population with yearly income between €15,000 and €26,000, and (v) the average estate and business income.
Finally, information on the number of RdC beneficiaries and the monetary amount of benefit received by the municipality was provided by the Italian Social Security Agency (INPS), which oversees the implementation of the scheme. The dataset identifies the total number of households and individuals receiving the scheme as of December 2019. Computing the share of RdC beneficiaries over the total municipal population, we obtained a mean value of 2.8%. However, this metric shows a very strong geographical distribution. As illustrated in Figure 3.3.1 below, municipalities in the southern regions have the highest share of RdC beneficiaries, reaching a maximum of 19.7%, while municipalities in the north and northeast present much lower shares. Nevertheless, municipalities with higher shares of RdC beneficiaries were observed even in the northern and central regions, suggesting that a disaggregated level of analysis is essential to evaluate the targeting of the policy.
To apply the SAE methodologies presented in the next section to the data gathered from the four sources mentioned above, all data-sets were aggregated across the three degrees of urbanisation areas in each of the 20 regions using Eurostat taxonomy applied to each munici-

pality. This classification clusters the European Local Administrative Units 2 (municipalities in Italy) by their population density (cities, towns and suburbs, rural areas). This originates a detailed grid of the European territory at the subregional level, disaggregating the regional level (NUTS2 level according to the 'Nomenclature of territorial units for statistics' used by Eurostat). This study considers each of the three DEGURBA categories across the 20 Italian regions as the main unit of analysis, for a total of 59 areas.

It is worth noting that owing to data provision constraints, the four datasets do not have the same exact temporal reference. HBS and Italian Ministry of Treasure data refer to 2017, EU-SILC data on poverty refer to 2017 (income refers to 2016), and the and INPS data refer to 2019. However, we believe that this slight temporal misalignment is irrelevant for two main reasons. First, absolute and relative poverty measures referring to a given area tend to be rather constant when considering a short run of one or two years (ISTAT, 2019). Second, even if data on RdC beneficiaries refer to 2019, the data used by INPS for allocating the benefit refer to the previous year, 2018, reducing the time lag reference of INPS data with the AROP and AP estimates to just one and two years, respectively.

Figure 3.3.1: Share of RdC beneficiaries in December 2019 by municipality.



3.3.2 Poverty definitions

The analysis builds on two definitions of poverty to estimate and define RdC's targeting performance. The first is absolute poverty. This indicator is based on yearly consumption data collected by the ISTAT through the HBS. Building a poverty indicator on consumption data, instead of current income information, provides a better proxy for permanent income. Consumption behaviours present less fluctuations in the short run compared to current income, thus avoiding the misclassification of households hit by temporary income shocks (Meyer and Sullivan, 2003). Denoting c_{ij} as the monthly household consumption of household j living in area i ($i = 1, \dots, m$), as t_{ij} the household-specific poverty line, N_i as the number of households living in area i , $I(u \leq k)$ as the indicator function (equal to 1 when $u \leq k$ and 0 otherwise), the absolute poverty (AP) indicator for an area i can be defined as:

$$AP_i = \frac{1}{N_i} \sum_{j=1}^{N_i} \mathrm{I}(c_{ij} \leq t_{ij}). \tag{3.1}$$

The absolute poverty threshold t_{ij} computed by ISTAT represents the monetary value, at current prices, of the basket of goods and services considered essential for each family, defined on the basis of the age of the members, the geographical distribution, and the type of municipality of residence. These parameters consider variations in the costs of living across the three main Italian main areas and across three different types of municipalities. Thus, the results are nine poverty lines indexed to local prices multiplied by the respective equivalence scale.

The second poverty indicator considered in this analysis is the At-Risk-of-Poverty Rate (AROP), an indicator developed by Eurostat which takes into consideration households' reported income based on information collected in the EU Statistics on Income and Living Condition Survey (EU-SILC). Denoting with y_{ij} the equivalised household disposable income of household j living in area i, and by t the poverty line, the AROP for an area i can be defined as:

$$AROP_i = \frac{1}{N_i} \sum_{j=1}^{N_i} \mathrm{I}(y_{ji} \leq t). \tag{3.2}$$

Therefore, this indicator classifies households as poor if their equivalised disposable income is less than poverty line t, which is usually set equal to 60% of the national median equivalent disposable income. That is, poverty is defined in relative terms, and a single national threshold is set for each country every year. Total household income is equivalised using the modified OECD equivalence scale, which assigns a weight of 1 to the first adult, 0.5 to any other adult members in the household, and 0.3 to children under the age of 14.

3.4 Methods

3.4.1 Small Area Estimation models

In this section, we describe the methods employed to obtain estimates at the region by DEGURBA level for the two indicators of poverty discussed previously. For both the AP and AROP estimates, our target indicators were the small-area means. The application of SAE models aims to increase the precision of direct survey estimates using administrative covariates at the DEGURBA-region level. The final aim of this analysis is to provide estimates for the assessment of policy targeting that could be replicated at different stages of the policy cycle. Following the guidelines set in a similar policy-oriented study commissioned by the Chilean Ministry for Social Development (Casas-Cordero Valencia et al., 2016), we apply relatively simple small-area estimation models which satisfy three main requirements:

1. the model should provide estimates for all 59 region-DEGURBA units in Italy;
2. the estimates should be close to the direct estimators for areas with large sample sizes;
3. the aggregated estimates for the areas should produce the official national estimate for the country.

For this purpose, we apply the bivariate Fay-Herriot (FH) model (Benavent and Morales, 2016), which is a multivariate version of the Fay-Herriot model (Fay III and Herriot, 1979), previously discussed in Section 2.4.1. The FH model and its multivariate transformations are area-level models that link the direct estimates to area-level covariates. They are especially useful when access to individual-level data is not available and the auxiliary variables and direct estimators are only available at an aggregated level. In this study, we apply both models - the FH and its bivariate version - to the two poverty indicators considered–the AP and the AROP. In the remainder of this section, we briefly describe the bivariate model. Section 3.5 presents the results of the best models selected based on the diagnostics performed. The detailed Fay-Herriot results are available in the supplementary materials.

3.4.2 Bivariate Fay-Herriot model

The bivariate Fay–Herriot model is a special case of the multivariate Fay-Herriot model (Benavent and Morales, 2016), an area-level linear mixed model that can be used to estimate the domain means of two correlated target variables. Let $\theta_i = (\theta_{i1}, \theta_{i2})'$ be a vector of the two characteristics of interest in area i, with $i = 1, \dots, m$, and let $\hat{\theta}_i^{direct} = (\hat{\theta}_{i1}^{direct}, \hat{\theta}_{i2}^{direct})'$ be a vector of direct estimators of θ_i. The bivariate Fay–Herriot model is defined in two stages. The model assumes that θ_i is linearly related to the auxiliary variables $X_i = diag(x_{i1}, x_{i2})$ with p explanatory variables $x_{ij} = (x_{ij1}, \dots, x_{ijp})$ with $j = 1, 2$ and $i = 1, \dots, m$ through the linking model

$$\theta_i = X_i\beta + u_i$$

whit $u_i \sim^{ind} N(0, AI_2)$, $i = 1, \dots, m$ where $\beta = (\beta_1', \beta_2')$ is a vector of coefficients and β_j where $j = 1, 2$ and are column vectors of size p, and A is the variance of the random effect. The direct estimator follows the sampling model

$$\hat{\theta}_i^{direct} = \theta_i + e_i$$

where $i = 1, \dots, m$, the vectors $e_i \sim N(0, D_i)$ are independent, and D_i is a 2×2 covariance matrix V_{ei} of the sampling errors. The bivariate Fay-Herriot model can be rewritten as

$$\theta^{direct} = X\beta + u + e \tag{3.3}$$

$u \sim N(0, AI_{2m})$; $e \sim N(0, D)$, where $\theta^{direct} = col_{1 \le i \le m}(\theta^{direct}{}_i)$, $X = col_{1 \le i \le m}(X_i)$, $u = col_{1 \le i \le m}(u_i)$, $e = col_{1 \le i \le m}(e_i)$, $D = col_{1 \le i \le m}(D_i)$ and the random effects u are independent of the sampling errors e. col is a matrix operator stacked by columns.
Under 3.3, the mean vector and the covariance matrix of θ^{direct} are

$$E(\theta^{direct}) = X\beta,$$

$$Var(\hat{\theta}^{direct}) = \Sigma = \Sigma(A) = AI_{2m} + D.$$

When the regression variance A is known, the true area mean is estimated by the best linear unbiased prediction (BLUP). In practice, A is unknown but can be estimated. Using the profile maximum likelihood, we can obtain the empirical BLUP (EBLUP)

$$\hat{\theta}^{BFH} = X\hat{\beta} + \hat{A}\hat{\Sigma}^{-1}(\hat{\theta}^{direct} - X\hat{\beta}) \tag{3.4}$$

and $\hat{\beta} = \hat{\beta}(\hat{A}) = (x'\Sigma^{-1}X)^{-1}X'\Sigma\hat{\theta}^{direct}$ and $\hat{\Sigma} = \hat{\Sigma}(A)$.

3.4.3 Targeting performance of anti-poverty measure

In the study of anti-poverty programs, the concept of targeting refers to the attempt by public officials to identify who is poor and then to restrict transfers to those individuals (see Hanna and Olken, 2018). Ravallion (2009) provides a useful overview of the four most commonly used indicators in the literature, focusing on assessing the targeted performance of anti-poverty measures.

Data on RdC beneficiaries are available for this research at the municipal level. This level of aggregation does not allow for the identification of recipients of RdC, who can be considered as not poor. As such, the most meaningful targeting indicator to be applied in this analysis is the Coverage Rate (CR) metric (Coady et al., 2004a,b). By defining D_{ij} as an indicator variable that takes value 1 if household j living in area i is the beneficiary of the RdC and, as in paragraph 3.3.2, by c_{ij} and y_{ij} the unit consumption and income measure, respectively, and with t_{ij} and t the corresponding poverty lines, two CRs can be defined as following:

$$CR_{iAP} = \frac{\sum_{j=1}^{N_i} D_{ij} \cdot \mathrm{I}(c_{ij} \leq t_{ij})}{\sum_{j=1}^{N_i} \mathrm{I}(c_{ij} \leq t_{ij})}$$

and

$$CR_{iAROP} = \frac{\sum_{j=1}^{N_i} D_{ij} \cdot \mathrm{I}(y_{ij} \leq t)}{\sum_{j=1}^{N_i} \mathrm{I}(y_{ij} \leq t)}.$$

The two measures above correspond to the ratio between the total number of households in absolute poverty and at risk of poverty who received the RdC living in area i, over the corresponding total number of households in absolute and relative poverty in area i. Given the aggregate nature of the data on RdC beneficiaries, in this application we make the assumption that $\sum_{j=1}^{N_i} D_{ij} \cdot \mathrm{I}(c_{ij} \leq t_{ij}) = \sum_{j=1}^{N_i} D_{ij}$ and $\sum_{j=1}^{N_i} D_{ij} \cdot \mathrm{I}(y_{ij} \leq t) = \sum_{j=1}^{N_i} D_{ij}$, defining therefore the following targeting measures:

$$CR_{iAP} = \frac{\sum_{j=1}^{N_i} D_{ij}}{\sum_{j=1}^{N_i} \mathrm{I}(c_{ij} \leq t_{ij})} \tag{3.5}$$

and

$$CR_{iAROP} = \frac{\sum_{j=1}^{N_i} D_{ij}}{\sum_{j=1}^{N_i} \mathrm{I}(y_{ij} \leq t)}. \tag{3.6}$$

3.5 Results and Discussion

3.5.1 SAE estimates of poverty

The following subsection presents the findings of the analysis related to the AP and AROP estimates through the application of the bivariate FH model defined above to the survey data previously described. All estimates, uncertainty intervals, and diagnostic metrics were obtained using the *emdi* (Kreutzmann et al., 2019) and *msae* (Permatasari and Ubaidillah, 2021) packages in *R*, version 3.6.0.

Model selection and diagnostics

Before presenting the findings of the SAE estimates on absolute and relative poverty across the 59 degree-of-urbanisation areas, we briefly discuss the model selection procedure. As discussed in Section 3.2, municipality-level covariates from the tax return data were aggregated at the target area level of analysis. The covariates employed for estimating the two target parameters were selected using a stepwise selection procedure. The final model for the estimation of AP and AROP showed correlation values between the bivariate FH estimates and the direct estimates of 0.76 for the AP and 0.93 for the AROP.

Gains in precision

Small area estimations were employed to improve the precision of direct estimates from both HBS and EU-SILC surveys designed to provide reliable information at higher geographical levels. To assess gains in the accuracy of our estimates, we compared the coefficient of variation of the bivariate FH model with those of the respective direct estimates. In this analysis, the application of SAE methods brought considerable gains to the precision of the estimates, as illustrated in 3.5.1. The bivariate FH model reduces the number of areas with CV estimates above the 33.3% threshold by more than three times compared to the direct estimates of absolute poverty, leaving only seven areas with an uncertainty of estimation too high to be considered reliable. In contrast, the bivariate FH estimates of the AROP show CVs below the 16.5% threshold. There are three main reasons for the difference in precision between the two estimates. First, the sample size of the EU-SILC data was significantly larger than that of the HBS data. This is reflected in the higher precision of the direct estimates of the AROP than those for absolute poverty. Second, the social phenomenon captured by the absolute poverty indicator is significantly rarer than that captured by the AROP. According to ISTAT, 6.7% of Italian households were considered living in absolute poverty in 2017, against the 20% considered at risk of poverty. As the variance of estimation is a function of the estimation itself, we expect larger uncertainty when the estimation focuses on rare events (Wolter, 2007). Finally, the differences in the definitions of these indicators, as discussed in Section 2, indicate that the model fits vary significantly. The covariates available for this analysis and included in both models almost entirely relate to income characteristics of the local areas and show a much stronger correlation with the AROP direct estimates than with the absolute poverty estimates.

Table 3.5.1: Comparison of the coefficients of variation of absolute poverty and AROP estimates.

		<16.5%	16.5-33.3%	>33.3%
AP	Direct	5	32	22
	FH bivariate	11	41	7
AROP	Direct	33	24	2
	FH bivariate	59	0	0

The distributions of confidence interval lengths are shown in 3.5.1 and Figure 3.5.2. The confidence interval lengths are reduced by applying the SAE models compared to direct estimation in both applications. These lengths decreased for all areas, confirming the necessity for the application of the SAE methods. Furthermore, the estimates obtained by applying the bivariate Fay and Herriot model always fall within the confidence interval of the direct estimates, with the exception of a single area, in support of the unbiasedness of the small area estimates.

Figure 3.5.1: Confidence intervals for the absolute poverty: model-based CIs are represented in black, CIs based on direct estimates are represented in red. Areas on the x-axis are ordered in ascending order of the direct estimate of AP.

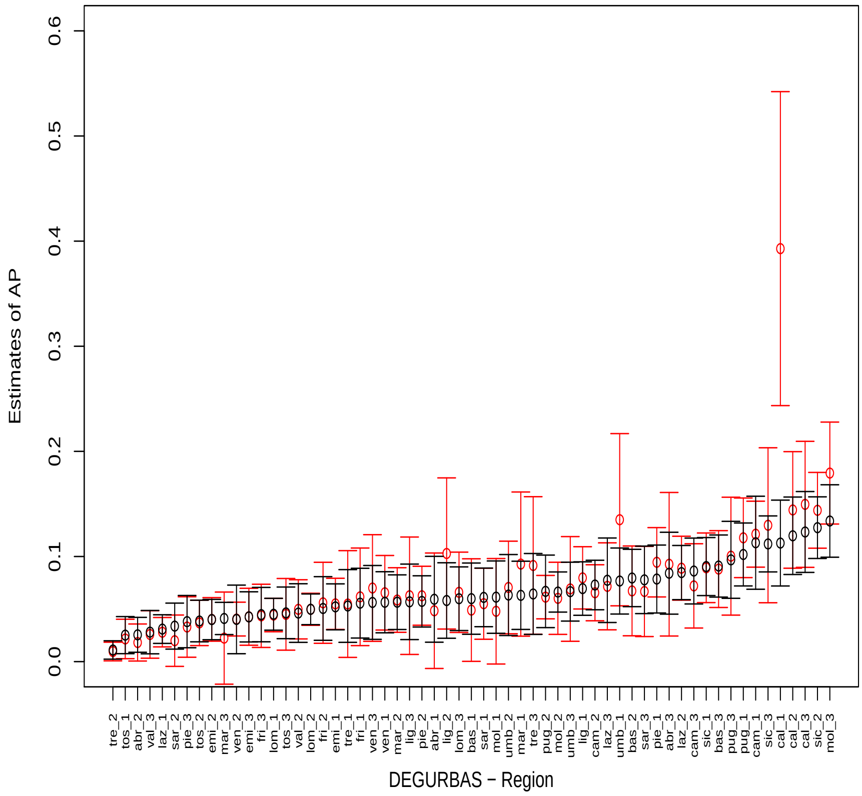
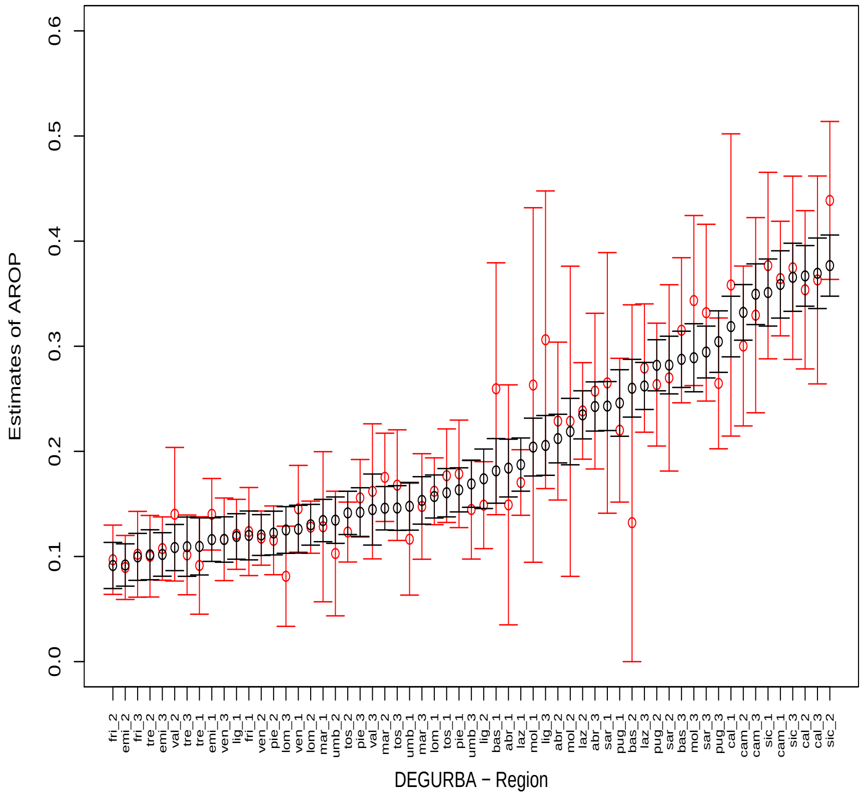


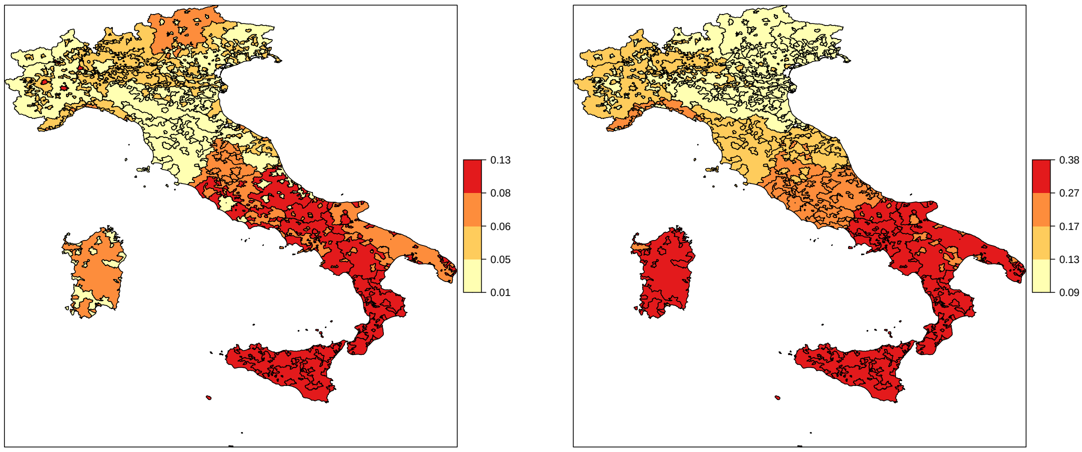
Figure 3.5.2: Confidence intervals for the AROP: model-based CIs are represented in black, CIs based on direct estimates are represented in red. Areas on the x-axis are ordered in ascending order of the direct estimate of AROP.



Discussion on poverty estimates at small area level

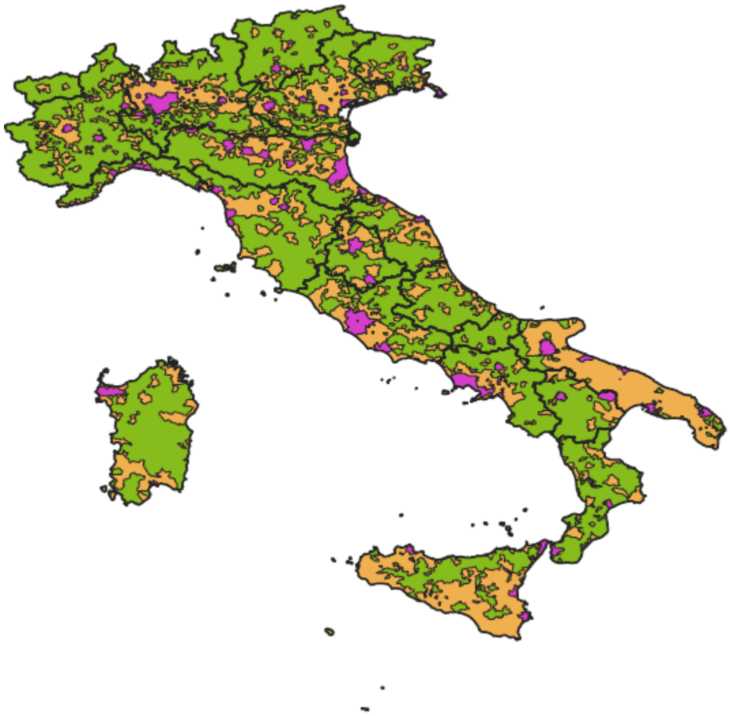
Figure 3.5.3 shows the distribution of poverty for 59 degrees of urbanisation across 20 Italian regions for both the AP and AROP indicators. To better understand the subdivision of Italian territory, Figure 3.5.4 represents the 59 areas and their degree of urbanisation.

Figure 3.5.3: Estimates of AP (left panel) and of the AROP (right panel) for the 59 degrees of urbanisation across 20 regions in Italy.



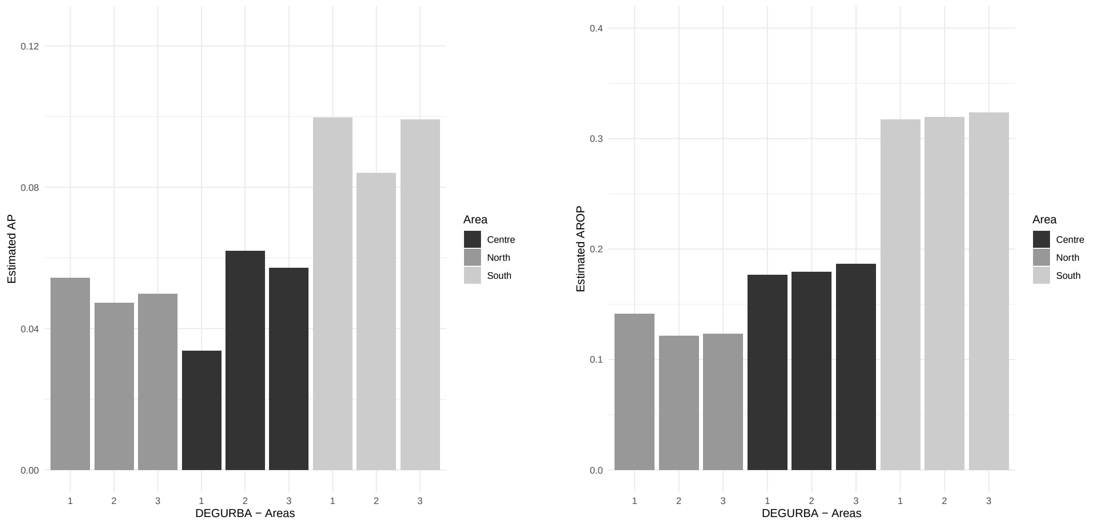
As we can see, both maps in Figure 3.5.3 show a clear distinction in poverty incidence across the country's three main areas of north, centre, and south: a higher poverty incidence characterises the southern areas. The AP index ranges from a maximum of 13.38% for the rural areas of Molise (South) to 1.11% in the suburban areas of Trentino-Alto Adige (North). The AROP ranges from 37.62% in suburban Sicilian areas (south) to 9.13% in suburban Friuli-Venezia Giulia (north). These findings reflect a country long-lasting economic dualism.

Figure 3.5.4: Region by DEGURBA classification of Italy: cities - pink, towns and suburbs - orange, rural areas - green, with regions represented with black boundaries.



An interesting trend is highlighted in the distribution of poverty. As illustrated in the left panel of Figure 3.5.5, absolute poverty estimates show how, both in the north and south of the country, urban areas (DEGURBA=1) have the highest values, followed by rural areas (DEGURBA=3). In the centre, urban areas show the lowest absolute poverty incidence, while suburban areas (DEGURBA=2) show the highest.

Figure 3.5.5: Estimates of AP (left panel) and of the AROP (right panel) for the degrees of urbanisation across Italy's three main areas (North, Centre, South).



Estimates of AROP, on the contrary, show a more heterogeneous incidence of poverty within the three main areas. In the centre and south, rural areas show the highest values, while the north urban centres have the lowest relative income. These differences emerge clearly in a careful comparison of the two poverty maps in Figure 3.5.3. First, following a clear north-south divide, the geographical distribution of absolute poverty incidence shows variation within the three main geographical areas. In the northern regions, we notice a higher poverty incidence, especially in the northeast of the country, with numerous areas in the second quartile of the absolute poverty distribution, compared to what is shown in the right-hand panel. Furthermore, we observe a high incidence of absolute poverty in the Piemonte-urban areas (approximately 9%). The incidence of absolute poverty in central areas appears to be significantly lower than that in the AROP indicator, with marked differences, especially in Toscana, Marche, and Lazio. Certain areas in the south show relatively lower levels of absolute poverty incidence compared to the AROP indicator, such as suburban areas in Campania, Puglia, and Sardinia.

The second main consideration is related to within-region heterogeneity in the incidence of absolute poverty, in contrast to the rather homogeneous within-region distribution of the AROP indicator. When poverty is measured by consumption, there seems to be greater variation within the same region across different degrees of urbanisation. In Lazio, we observe a statistically significant difference between absolute poverty in urban and both rural and suburban areas, with the former showing a poverty incidence within the bottom quartile of the national distribution, while the latter in the top two quartiles. Similar statistically significant differences were observed in other areas and regions, such as Molise and Abruzzo. The considerations highlighted thus far are the result of differences in the definition of poverty indicators considered in the analysis. As discussed in Section 2, absolute poverty is estimated on the basis of consumption behaviour based on different poverty lines, varying across Italy's three main areas and across the size and type of the municipality of the survey respondent. Unlike single national poverty thresholds, such as the AROP indicator present in the EU SILC data, this approach allows us to capture differences in the costs of living. In a context with marked geographical heterogeneity, capturing households' real purchasing power is essential to gain an accurate picture of the incidence of poverty. Indeed, there is clear evidence of variation in price levels across Italian regions. A study by ISTAT (2010) highlighted how Purchasing Power Parities (PPP) were heterogeneous among the 20 regional capitals, with Bolzano (Trentino-Alto Adige) recording the highest costs of living (PPP = 105,5, with Italy = 100) and Napoli (Campania) as the town where the cost of living was the lowest (PPP = 93,8). An indicator, such as the absolute poverty metrics developed on the HBS, attempts to account for these differences, moving away from the traditional single national poverty thresholds. The difference in these approaches is evident, and our application provides a valid example. The urban areas in Lazio and Piedmont, comprising of two of the largest cities in Italy, Rome, and Turin, present very different AROP. Turin is in the bottom quartile of the national distribution, whereas Rome is in the third quartile. Once the costs of living are considered and consumption data are utilised, Rome moves to the bottom quartile and Turin to the highest. The implications of using two different indicators of poverty for policy are also evident and will be discussed in the next section.

3.5.2 Discussion on RdC targeting

Following the discussion of the local area estimates of the AP and AROP, the analysis moves to an assessment of the RdC-targeting performance against such indicators. Given the limitations of the data on RdC beneficiaries, which are available at the aggregate municipality level, and the difficulty of excluding non-poor recipients from the overall share, the resulting targeting indicators will likely be an overestimation of the true parameter. Figure 3.5.6 plots the two CR indicators of the RdC for each of the 59 DEGURBA areas across the 20 Italian regions. Appendix 3.B provides the CR estimates as well as the SAE estimates of both AROP and AP for all areas. Once again, we observe a rather heterogeneous distribution of the CR indicators across the Italian territory. The main difference in the comparison of the two indicators is the width of the range of values. The CR_{AP} indicator ranges from 5.6% in the rural areas of Trentino-Alto Adige to 179.31% in the suburban areas of Sardinia. In contrast, the Values of the CR_{AROP} indicator show a significantly narrower range, from 3.31% of rural areas in Trentino-Alto Adige to 31.07% of urban areas in Sicily. The CR_{AROP} indicator highlights how the vast majority of households identified as at risk of poverty are excluded from the support provided by the RdC. On the contrary, the CR_{AP} indicator describes a policy with large geographical heterogeneity in its targeting performance, excluding a large number of absolute poor households in areas with higher costs of living, and including non-poor households in more affordable ones. Overall, the policy seems to consistently show lower targeting performance in the northern areas of the country, especially in the north-east. Approximately all of the bottom 10 areas for both the CR_{AP} and CR_{AROP} indicators are in the north. Moreover, if we consider both the CR_{AP} and CR_{AROP} metrics, we find that rural areas across Italy present lower targeting performance, irrespective of the three main areas considered, as illustrated in Figure 3.5.7. We obtain similar results if we consider measures of absolute poverty using other SAE methods as shown in the Appendix in Section 3.B in Figure 3.B.1 and Figure 3.B.2.

This may depend on that fact that households living in rural areas might face higher information barriers to accessing the program, such as limited access to digital services and in-person support (Shucksmith, 2003), or might be subject to stronger social stigma related to participating in welfare programs (Currie, 2004). Further research on the reasons behind this trend could guide future policy interventions. These findings add new and important evidence to the literature on this topic. First, if we consider studies assessing the overall targeting of RdC at the national level (Baldini et al., 2019; Curci et al., 2020), our analysis reveals important geographical heterogeneity in targeting performance by focusing on sub-regional areas. Second, in relation to other analyses focusing on the geographical distribution of RdC beneficiaries (Checchi et al., 2021), we move beyond the correlation of RdC distribution with geographical socioeconomic variables to provide punctual estimates for each sub-regional area on the extent to which the measure succeeds in reaching its anti-poverty objective.

Figure 3.5.6: Coverage rate of RdC estimated on AP (CR_{AP} - left panel) and on AROP (CR_{AROP} - right panel) for the 59 degrees of urbanisation across Italy's 20 regions.

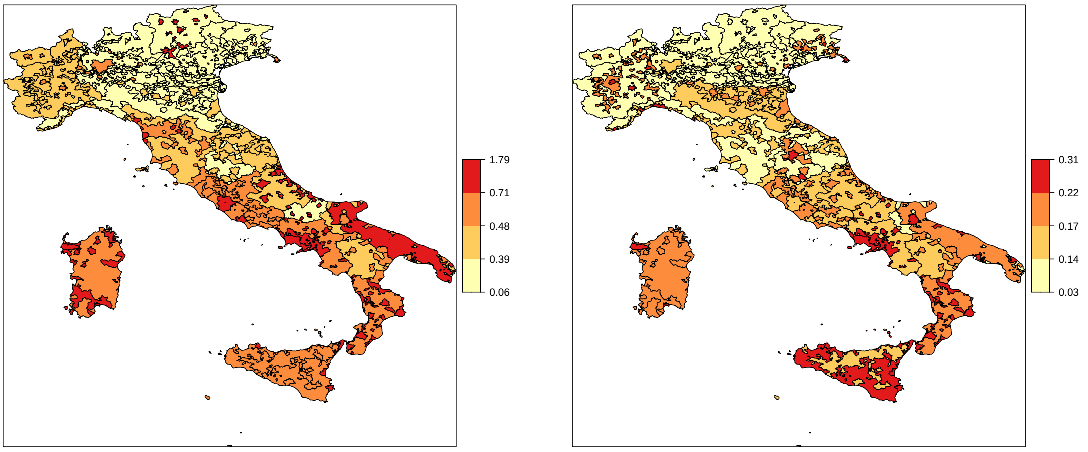
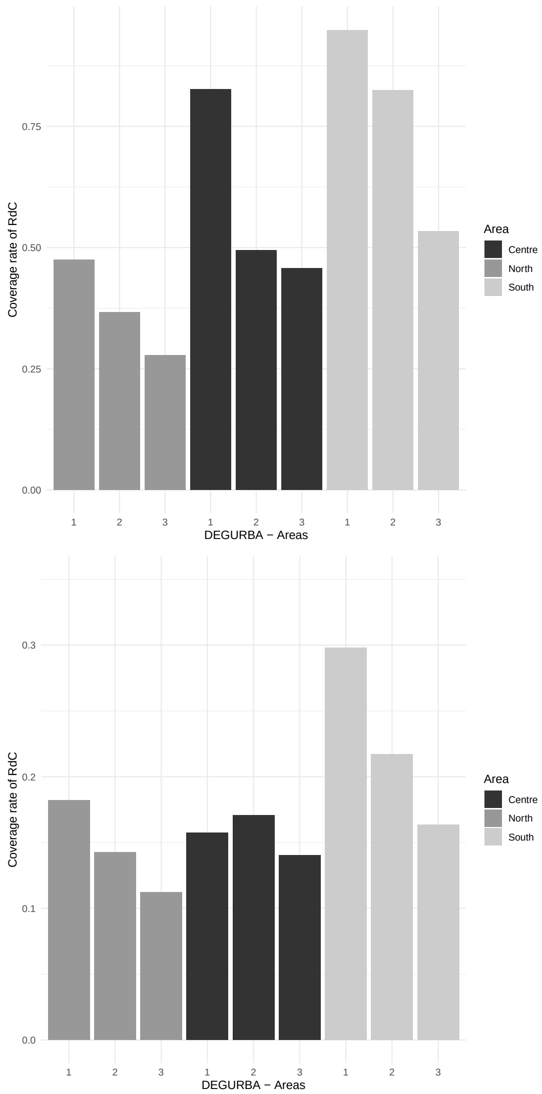


Figure 3.5.7: Coverage rate of RdC estimated on absolute poverty (left panel) and on AROP (right panel) for the three degrees of urbanisation in Italy across main areas.



A number of potential explanations are at play for understanding the heterogeneous targeting performance of the policy. As discussed in 3.1, scholars and experts have highlighted how specific aspects related to RdC design might undermine the targeting efficacy of the method. These are the 10-year residence requirement for non-EU citizens, the choice of an $adhoc$ equivalence scale which penalises larger families, and the high marginal tax rate, which increases the costs of claiming in relation to the opportunity of employment.
In order to test whether these factors might be associated with the high variations in CR_{AP} values, we consider a multivariate Ordinary Least Squares (OLS) model. The dependent

variable is the log-transformed CR_{AP} score, while the independent variables are the average family size, percentage of foreign residents, and level of employment in all 59 areas, each capturing one of the three potential characteristics of the policy design affecting its targeting performance. Table 3.5.2 presents the OLS results. Employment has a significant and negative coefficient. This finding is robust to a series of further checks. First, when the sample is restricted to those areas with absolute poverty CV estimates below the 33.3% threshold (Model (2) in Table 3.5.2) the negative coefficient persists. Second, we observe similar results if we include macro-area fixed effects in the model, capturing in this way any unobserved characteristics related to Italy's three main geographical areas which might affect the RdC targeting performance (Model (3)). For example, regions in the North have larger per capita spending on welfare programs which might represent an alternative to claiming RdC. Finally, the negative coefficient of the employment variable is also observed when considering models including the CR indicator obtained through other SAE methods, such as a univariate and arcsin-transformed FH model (see Section 3.B).

The literature suggests two potential mechanisms underlying these findings. The first identifies low targeting performance as the non-take-up of benefits by eligible households. This is due to the high opportunity costs of claiming in relation to employment opportunities (Currie, 2004). Eligible claimants have low incentives to claim if they know that they can find employment easily. These incentives are made even lower by the prospect of being subject to strict conditionality regimes attached to program participation and by the high marginal tax rate of the benefit. The second mechanism linking higher employment rates to lower targeting performance is related to the demographics of poverty and, in particular, to the rising phenomenon of the so-called "working poor". The aftermath of the 2008 financial crisis witnessed an increasing trend in work poverty across Europe and Italy (Saraceno, 2015). In 2107, an estimated 12% of employed households lived under the relative poverty threshold (Saraceno, 2020). The 2017 HBS indicates that 55.6% of all households in the absolute poverty report were employed. This percentage rises to 68.4% in northern areas, where the CR based on absolute poverty estimates is significantly lower. While the phenomenon of the working poor is captured in official surveys such as the EU-SILC and HBS, specific aspects of the RdC design are likely to exclude the working poor households from the program. First, working poor live, on average, in larger households than other households in poverty, and are therefore disadvantaged by the relatively less generous equivalence scale adopted in the RdC. The 2017 HBS, indeed, indicates the average size of poor households with at least one member in employment, to be 3.36 members against 2.21 for other households in poverty.

Second, working poor are likely to hold higher savings and financial capital; thus, the current RdC saving threshold might exclude them from the program. These findings offer clear evidence of how an anti-poverty program such as the RdC is ineffective in targeting poverty in rural areas and in areas where rates of in-work poverty are higher.

Table 3.5.2: OLS results

VARIABLES	(1)	(2)	(3)
Avg. family size	-0.427	-0.549	-0.601
	(0.404)	(0.416)	(0.409)
% Foreign residents	3.142	3.006	6.318**
	(2.217)	(2.314)	(2.891)
% Employment	-7.445***	-8.221***	-5.074***
	(1.218)	(1.239)	(1.818)
Area = North			-0.204
			(0.163)
Area = South			0.464*
			(0.266)
Constant	6.420***	7.337***	4.394**
	(1.523)	(1.533)	(2.013)
Observations	59	52	52
R-squared	0.446	0.526	0.575
Shapiro-Wilk test p-value	0.020	0.053	0.076

Notes: Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. The dependent variable is the log transformed CR indicator. *Avg. family size* describes the average size of families in each area. *% Foreign residents* is the share of foreign residents in each area. *% Employment* is the share of employment in each area. Column (1) presents the OLS results across all 59 areas. Column (2) shows the OLS result considering only the 52 areas where absolute poverty estimates CVs are below the 33.3% threshold. Column (3) includes in the model the fixed effects for each of Italy's three main areas, with "Centre" as the reference category.

3.6 Conclusions

To the best of our knowledge, we presented the first study on the targeting of the "Reddito di Cittadinanza" (RdC) anti-poverty policy at the local level. The study was based on four main data sources (including surveys and administrative data) and made use of appropriate statistical methods (such as small area estimation techniques) to obtain reliable poverty estimates for the 59 local areas of interest. The Italian regions were further classified according to the degree of urbanisation. It is essential to implement local-level targeting of anti-poverty policies to meet the needs and problems of the territory where people live and to develop a successful policy program. The results of this study show a heterogeneous targeting performance of the RdC policy, with a generally lower targeting affecting northern regions and rural areas. Overall, the analysis highlights how an anti-poverty measure, such as the RdC, has limited coverage if we consider the CR_{AROP} indicator, reaching, on average, only 17.9% of households in relative poverty. However, once the CR_{AP} indicator is considered, the findings lead to different conclusions. Indeed, the policy fails to consider subnational differences in the costs of living. The result is a measure of a small proportion of poor households in areas where costs of living are higher, while providing support to large shares of households that do not fall under the definition of absolute poverty in more affordable areas. These findings have different implications from a policy perspective. First, to improve overall targeting, more funding should be allocated to the measure to expand the cohort of beneficiaries. The findings point to the working poor as a group largely excluded from this measure. This group could benefit from a more generous equivalence scale for larger families and from a higher savings threshold. Second, policymakers should consider adjusting the support available to subnational variations in the costs of living. For example, the amount of housing support available could be linked to regional rent and property prices, as in other income support programs (e.g., Universal Credit in the UK or Hartz IV in Germany). Third, more effort is required to understand the drivers behind low take-up in rural areas, ensuring that poor households have access to adequate information, digital infrastructure, and literacy to access support they are eligible for. Finally, the significant differences in targeting performance based on two different poverty indicators highlight the importance of setting clear definitions when formulating policy objectives and design to obtain clearer and more transparent evaluation $expost$.

From a methodological point of view, it is important to emphasise that the poverty estimates used to evaluate the targeting, computed using a bivariate Fay-Herriot small area model, are robust to different model specifications. Similar results were indeed obtained with a univariate version of the model, also considering an arcsin transformation of the FH model for poverty indicators. Future developments of the present work should focus on estimating the variability of the targeting indices to better evaluate the significant differences affecting the Italian territory.

Appendix

3.A Bivariate FH AP estimates, Bivariate FH AROP estimates and CR estimates

	Domain	Region	Degurba	Area	AP_rate	AP_CV	CR_AP	AROP_rate	AROP_CV	CR_AROP
1	101	piemonte	1	nord	7.96	17.49	53.68	16.32	6.55	26.17
2	102	piemonte	2	nord	5.73	21.70	45.55	12.22	8.68	21.36
3	103	piemonte	3	nord	3.80	33.32	45.15	14.19	8.38	12.11
4	202	valle d'aosta	2	nord	4.64	27.01	50.16	10.84	10.32	21.47
5	203	valle d'aosta	3	nord	2.80	37.52	40.81	14.45	11.92	7.92
6	301	lombardia	1	nord	4.51	17.14	50.74	15.69	6.65	14.57
7	302	lombardia	2	nord	4.98	14.98	32.51	13.00	7.60	12.44
8	303	lombardia	3	nord	5.98	25.81	24.56	12.51	9.04	11.75
9	401	trent. a.d	1	nord	5.29	33.34	21.59	10.94	12.65	10.44
10	402	trent. a.d	2	nord	1.11	40.27	73.03	10.15	11.93	7.96
11	403	trent. a.d	3	nord	6.43	30.51	5.63	10.92	13.16	3.32
12	501	veneto	1	nord	5.65	26.27	38.31	12.58	9.21	17.19
13	502	veneto	2	nord	4.10	19.05	31.95	12.03	8.23	10.90
14	503	veneto	3	nord	5.63	31.84	21.22	11.60	9.48	10.30
15	601	fvg	1	nord	5.56	30.44	55.37	11.99	9.88	25.68
16	602	fvg	2	nord	5.05	30.57	33.16	9.13	12.24	18.35
17	603	fvg	3	nord	4.47	29.51	27.38	9.95	11.44	12.30
18	701	liguria	1	nord	6.95	18.67	45.29	11.89	9.27	26.46
19	702	liguria	2	nord	5.93	35.13	43.05	17.37	8.30	14.70
20	703	liguria	3	nord	5.69	32.21	37.78	20.54	7.05	10.46
21	801	emilia romagna	1	nord	5.21	21.41	41.91	11.59	9.13	18.83
22	802	emilia romagna	2	nord	4.00	24.44	37.28	9.18	11.17	16.25
23	803	emilia romagna	3	nord	4.26	28.60	34.26	10.18	10.37	14.32
24	901	toscana	1	centro	2.52	35.71	99.76	16.05	7.30	15.66
25	902	toscana	2	centro	3.86	26.15	59.56	14.13	7.42	16.28
26	903	toscana	3	centro	4.62	30.79	40.74	14.59	7.42	12.90
27	1001	umbria	1	centro	7.74	26.49	45.11	14.76	7.84	23.66
28	1002	umbria	2	centro	6.31	26.26	40.94	13.44	8.37	19.21
29	1003	umbria	3	centro	6.68	26.33	34.63	16.88	6.72	13.71
30	1101	marche	1	centro	6.33	31.03	39.31	13.40	7.65	18.58
31	1102	marche	2	centro	5.65	23.45	42.28	14.57	7.22	16.40
32	1103	marche	3	centro	4.02	41.46	47.26	15.32	7.51	12.39
33	1201	lazio	1	centro	3.09	22.47	91.34	18.72	6.89	15.10
34	1202	lazio	2	centro	8.47	15.51	48.42	23.44	4.97	17.50
35	1203	lazio	3	centro	7.67	20.90	52.72	26.19	4.35	15.43
36	1301	abruzzo	1	sud	5.81	31.51	74.23	18.39	7.62	23.46
37	1302	abruzzo	2	sud	2.55	33.23	157.63	21.19	5.56	18.96
38	1303	abruzzo	3	sud	8.41	23.63	39.67	24.24	4.92	13.76
39	1401	molise	1	sud	6.13	28.62	80.16	20.38	6.88	24.13
40	1402	molise	2	sud	6.66	21.46	70.88	21.85	7.37	21.59
41	1403	molise	3	sud	13.38	13.14	31.99	28.87	5.71	14.82
42	1501	campania	1	sud	11.20	12.14	96.63	35.84	4.55	30.19
43	1502	campania	2	sud	7.29	16.48	103.99	33.18	4.06	22.84
44	1503	campania	3	sud	8.62	18.55	61.23	34.91	4.21	15.12
45	1601	puglia	1	sud	10.19	14.99	69.25	24.57	6.56	28.72
46	1602	puglia	2	sud	6.62	14.71	79.82	28.15	4.40	18.77
47	1603	puglia	3	sud	9.68	19.29	40.00	30.40	4.90	12.74
48	1701	basilicata	1	sud	5.99	28.85	55.85	18.12	8.64	18.46
49	1702	basilicata	2	sud	7.78	21.06	70.37	25.97	5.39	21.07
50	1703	basilicata	3	sud	9.09	16.58	44.85	28.72	4.75	14.20
51	1801	calabria	1	sud	11.31	19.96	81.05	31.83	4.61	28.79
52	1802	calabria	2	sud	11.97	15.70	83.86	36.65	4.01	27.39
53	1803	calabria	3	sud	12.33	15.91	59.05	36.89	4.63	19.74
54	1901	sicilia	1	sud	9.04	15.57	120.51	35.06	4.64	31.08
55	1902	sicilia	2	sud	12.74	11.74	64.44	37.63	3.94	21.82
56	1903	sicilia	3	sud	11.28	18.48	51.21	36.51	4.52	15.82
57	2001	sardegna	1	sud	6.11	23.28	115.63	24.28	4.89	29.10
58	2002	sardegna	2	sud	3.38	33.06	179.31	28.18	4.96	21.50
59	2003	sardegna	3	sud	7.86	20.98	66.72	29.41	4.27	17.82

3.B AP and CR_{AP} estimates with univariate FH and arcsin-transformed FH models

Figure 3.B.1: Absolute poverty estimates obtained with a univariate FH model (left panel) and with arcsin-transformed FH model (right panel) for the 59 small areas in Italy.

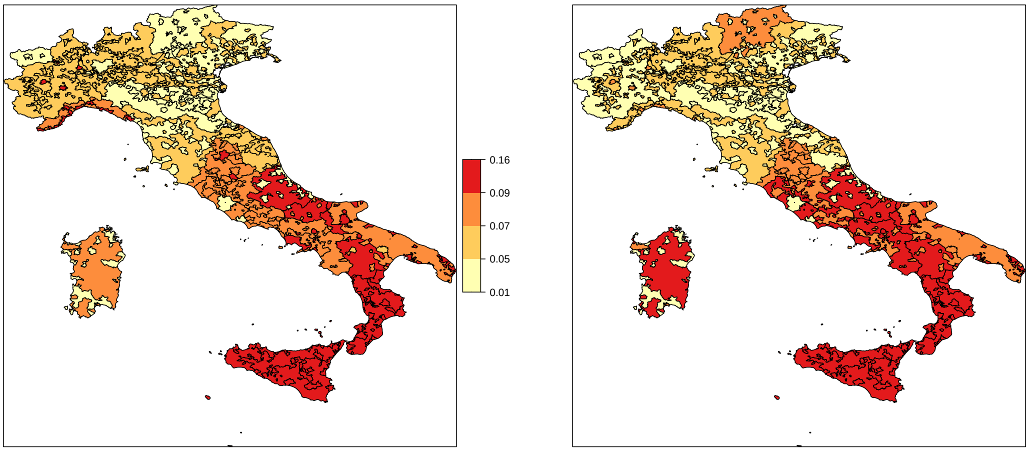
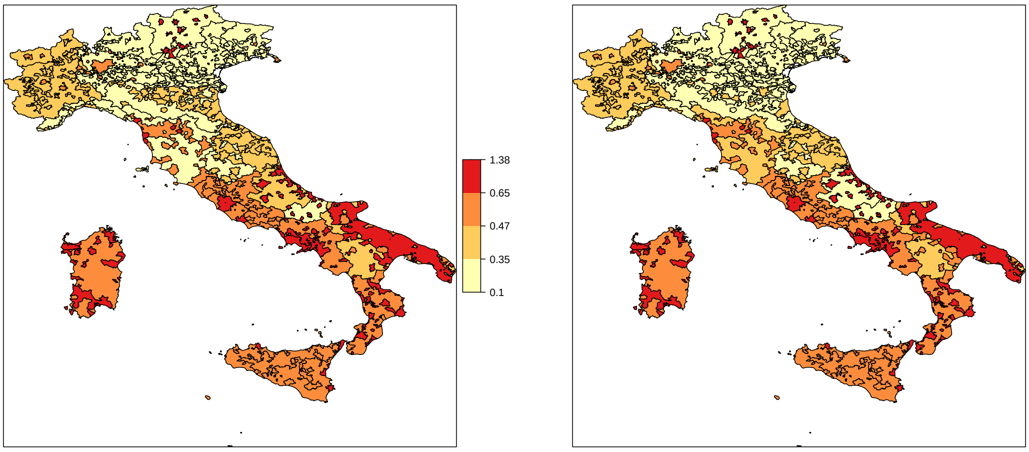


Figure 3.B.2: CR_{AP} estimates. AP obtained with a univariate FH model (left panel) and with arcsin-transformed FH model (right panel) for the 59 small areas in Italy.



3.C OLS with CR_{AP} estimates obtained with a univariate FH model and with arcsin-transformed FH model.

VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)
Avg. family size	-0.102	-0.154	-0.229	-0.436	-0.591	-0.662
	(0.350)	(0.295)	(0.308)	(0.394)	(0.404)	(0.403)
% Foreign residents	3.357*	3.453**	4.312**	3.229	2.562	4.930*
	(1.922)	(1.632)	(2.049)	(2.164)	(2.215)	(2.742)
% Employment	-7.023***	-7.399***	-6.108***	-7.480***	-8.164***	-5.513***
	(1.056)	(0.896)	(1.516)	(1.189)	(1.189)	(1.782)
Area = North			-0.0920			-0.198
			(0.126)			(0.160)
Area = South			0.163			0.350
			(0.218)			(0.256)
Constant	5.222***	5.588***	4.567***	6.405***	7.360***	5.006**
	(1.320)	(1.137)	(1.564)	(1.487)	(1.488)	(1.959)
Observations	59	53	53	59	54	54
R-squared	0.503	0.635	0.644	0.459	0.541	0.576
Shapiro-Wilk test p-value	0.201	0.664	0.496	0.022	0.080	0.100

$Notes$: Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. The dependent variable is the log transformed CR_{AP} indicator obtained with the arcsin transformed FH model (columns (1), (2), (3)) and wit the univariate FH model (columns (4), (5), (6)) . *Avg. family size* describes the average size of families in each area. *% Foreign residents* is the share of foreign residents in each area. *% Employment* is the share of employment in each area. Columns (1) and (4) present the OLS results across all 59 areas. Columns (2) and (5) show the OLS result considering only the areas where absolute poverty estimates CVs are below the 33.3% threshold. Column (3) and (6) includes in the model the fixed effects for each of Italy's three macro areas, with "Centre" as the reference category.

Chapter 4

The political economy of RdC: an assessment of the policy electoral impact

Abstract: We study the electoral impact of the Citizens' Income (RdC), a minimum income scheme introduced in Italy by a populist coalition government in 2019. At municipality level, a 1 pp increase in the share of recipients correlates with a 1.97 pp increase in the support for the Five Star Movement (M5S), the senior coalition partner and main proponent of the program. However, this positive correlation is not robust when we take into account that 5SM was already stronger in municipalities with more beneficiaries prior to the introduction of the RdC. Using a difference-in-difference strategy and leveraging quasi-exogenous variation in the share of beneficiaries, we find that RdC had opposite effects on the M5S performance in the two nationwide elections held after its introduction. At the 2019 European elections, it had a negative effect stemming from both the excessive expectations (and subsequent disappointment) it generated ahead of the 2018 general elections and due to its design that penalises large families. At the 2022 general elections, it had a positive effect in response to an electoral campaign to abolish the program by the main opposition (and front-runner) party. To better identify and contextualize this latter effect, we use as a counterfactual the M5S performance in the contemporaneous administrative election. The findings of this study support the argument that the politics behind cash transfers matter. In the case of a partisan policy like the RdC, voters respond to parties' stands and cast their preferences accordingly. Whether this is reflected in support for the incumbent, however, depends on the interacting combination of the type of elections and the extent to which

voters' expectations were met.

4.1 Introduction

A complete analysis of cash transfers requires both an assessment of their redistributive effects as well as their politically consequential aspects (Golden and Min, 2013). As such, following an analysis of the Citizens' Income (RdC) targeting, this chapter focuses on the political consequences of the measure, which are largely related to its policy design and implementation. As previously described in Chapter 1, the trend of economic decline accelerated since the 2007/2008 financial crisis not only led to the issue of poverty becoming a focal policy point with the introduction of the RdC, but has also catapulted it to the forefront of the political debate. In particular, the M5S, an anti-establishment political party that disrupted the Italian political arena starting in 2010, heralded the introduction of a guaranteed minimum income scheme as one of the key manifesto pledges right from the first national elections it took part to in 2013, to then implement it into policy as the senior party of a coalition government following the electoral success of the 2018 elections. Since its inception, the policy attracted strong criticisms from all ends of the political spectrum and strongly polarised the debate in relation to the government's role in providing support to low-income households. Such debate featured prominently in the electoral campaign ahead of the 2022 elections, where parties at the centre and to the right of the spectrum pledged to overhaul the program.

In this context, the study of how the introduction of the RdC impacted the electoral support for the main political force which campaigned for and later enacted its introduction provides useful insight beyond voters' responses to cash transfers, shedding light on how electoral pledges affect their voting behaviours. A vast body of literature has investigated the impact of conditional cash transfers, such as the RdC, on the electoral support of incumbent parties. The emerging consensus, is these programs improve incumbents' electoral performance (Zucco Jr, 2013; De La O, 2013; Manacorda et al., 2011) These studies have typically focused on a measure implemented by the incumbents. In the case of the RdC, on the contrary, we assess how an electoral promise of a large welfare program, subsequently enacted into policy, translates into a party's electoral victory and future performance in the next elections. Gromadzki et al. (2022) investigate a similar dynamic in the context of Poland, where a populist party first campaigned and later introduced a universal child benefit program. Contrary to their case, however, the RdC does not constitute a universal type of support program, allowing us to investigate how mismatches between electoral promises and enacted eligibility criteria impact voters' behaviours. Finally, the partisan nature of the RdC makes this research an important case study in the analysis of how the politics behind the implementation of a policy shape its political consequences (Imai et al., 2020).

In this chapter, we build on a panel of municipality-level election data to study the electoral impact of the RdC. Even controlling for a number of municipality characteristics, we show that the intensity of the policy measure is positively associated with the M5S electoral outcome: a 1pp increase in the share of recipients correlates with a 1.97pp increase in the support for the M5S. However, this positive correlation is not robust when we take into account that M5S was already stronger in municipalities with more beneficiaries. Using a difference-in-difference strategy and leveraging quasi-exogenous variations in the share of beneficiaries, we find that the RdC had opposite effects on the M5S performance in the two national-wide elections held after its introduction. In the 2019 European elections, it had a negative effect stemming from both the excessive expectations (and subsequent disappointment) the policy proposal had generated ahead of the 2018 General elections and from aspects related to its policy design which penalizes large families. In the 2022 general elections, on the contrary, we find a positive effect, in response to a campaign against the program by the main opposition (and front-runner) party.

The structure of the chapter is as follows. In the next section, we provide an overview of the RdC key features and a timeline of events that led to its implementation as well as to the political evolution of the M5S. In Section 4.3 we discuss the conceptual framework behind some of the most relevant theories related to voters' responses to redistributive policies. Section 4.4 presents the data employed in the analysis and lays out some of the descriptive findings, while Section 4.5 introduces the empirical strategy adopted. Section 4.6 presents the main findings and proposed mechanisms, before concluding in Section 4.7.

The research findings presented in this chapter are relevant in at least two ways. First, they provide a detailed account of the process of political polarization of the issue of poverty in the context of Italian politics, by trying to disentangle the concomitance of factors associated with the establishment of the M5S as the most representative political force in the most economically disadvantaged areas of the country. Second, they contribute to the literature on the political consequences of redistributive policies by providing, to the best of our knowledge, a first and unique assessment of how voters respond to both electoral pledges and the actual implementation of a highly partisan cash-transfer program.

4.2 Institutional background

4.2.1 RdC key features

As presented in Section 3.2.1, the RdC is the largest minimum income scheme ever approved in Italy (Baldini and Gori, 2019). The program's total monetary transfer was €3.9 (€3.7 without its component that goes to people that have passed the retirement age and known as Citizens' Pension (PdC)) billion in 2019 (from April to December), it peaked at €8.8 (€8.4 without PdC) billion in 2021, and decreased to €7.9 (€7,6 without PdC) in 2022.
In the context of this chapter, it is worth reiterating the eligibility criteria established by the policy. First, households' Equivalent Economic Situation Indicator (ISEE), an index that reflects both annual income and wealth, needs to be lower than €9,360. Second, they cannot own a real estate property that exceeds €30,000 (excluding its main residence) nor possess certain kinds of vehicles or boats and crafts. Third, their financial assets must be lower than €6,000 for single-person households and up to €10,000 for bigger ones. Fourth, their yearly equivalent income must not exceed €6,000 (€9,360 for households living in rented accommodation). As previously discussed, the equivalence scale employed in the scheme is set to penalise large families as it assigns a value of 1 to the household head, 0.4 to each additional adult member, and 0.2 to each child with an upper bound of 2.1. As we shall explain more in detail below, this was the result of the electoral commitment of bringing up all household incomes (including singles) to €780. Finally, the claimant must have either EU citizenship or hold an EU residence permit and reside in Italy for at least ten years.

4.2.2 RdC politics and implementation

A program named RdC appeared for the first time as a policy proposal in a M5S official document in January 2013 two weeks ahead of the 2013 national elections (see Figure4.2.1). In an open letter to voters, Beppe Grillo, the movement co-founder, and most prominent leader, listed RdC as his first policy prescription to "rescue Italy". Although the letter did not go into much detail, the name chosen (the English translation of RdC is Citizens' Income) suggests that its spirit was close to a universal basic income measure. Hence, it did not require to specify eligibility rules nor to choose a poverty line.
The second version of the RdC is contained in a parliamentary bill presented by M5S in October 2013. Despite its name, it was not designed as a universal basic income but as a targeted minimum income scheme. The proposal was quite generous and contained broad eligibility requirements: it envisaged a monthly benefit addressed to all households living in relative

poverty, defined as those whose income was below the 60% of the national median. In Italy, such a threshold corresponded to a monthly equivalent income of €780. As we shall see in detail below, this number affected the design of the third and final version of the RdC and, possibly, anchored the expectations of potential beneficiaries to a total benefit that is relatively high for single household recipients. Note that the equivalence index used in the bill was the so-called OECD-modified scale (that assigns a value of 1 to the household head, 0.5 to each additional adult member, and 0.3 to each child) and the eligibility conditions did not make any reference to wealth indicators. As a result, the pool of potential beneficiaries comprised an estimated 5 million families and more than 10 million people, with an estimated overall cost of about €17 billion[1]. More generally, the emphasis of the proposal was on employment and the risk of poverty, rather than solely poverty. In particular, it comprised an important work activation program justified by the assumption that "the lack of employment opportunities is the main cause of poverty", as stated in the official bill's documentation.

The third (and finally implemented) version of the RdC was approved in January 2019, after the sweeping victory obtained by the M5S at the March 2018 national election and before the May 2019 European elections. A number of features set aside the enacted measure from its original policy proposal. First, its estimated cost (around €8 billion) falls a long way short with respect to the estimated cost of the previous version (around €17 billion or more, as mentioned above); second, eligibility criteria include both income and wealth conditions and this restricts the pool of beneficiaries. On the other hand, the M5S communication strategy remained the same. The policy was presented officially as "a revolution for the labor market" even if its scope in terms of employment creation was quite limited. Moreover, the equivalence scale applied to households in order to compute the amount of the benefit was biased by the promise to give €780 (the poverty threshold mentioned above) to every person in poverty. As a result, the program turns out to be generous for singles and couples, while it is rather tight for families with more than two children. Poor communication strategy and bad design are two key features of the RdC and both could have backfired in terms of actual electoral support for the measure, especially in areas with high unemployment and demography characterized by larger families.

Two factors can explain why the M5S overlooked these flaws. First, the party was in a rush to approve and implement the RdC bill to reap possible electoral dividends before the 2019

[1]This estimate, calculated by the National Bureau of Statistics (ISTAT), may underestimate the true cost of the policy. According to the National Social Security Institute (INPS) the overall cost could have been around €30 billion.

European elections. Hence, poor communication and bad design may be related to time constraints imposed by the partisan nature of the measure. Second, the RdC elaboration and approval are prototypical cases of populist policymaking. M5S decided to centralize the design of the final scheme of the bill, hold a simplistic view of its problems, and refused to engage with experts and policy advisors (Gori, 2020).
The partisan nature of the policy also meant that, right from its implementation, the program was met with strong criticism from all ends of the political spectrum. While some of the critiques from centre-left political formations mostly centered around issues related to its policy design, the most vocal push-backs to the program came from centre and right-wing parties. Based largely on anecdotal evidence of fraud as reported in mainstream media and on the low success of the program active labour market component, the centrist party Italia Viva and the far-right party Brothers of Italy (FdI) mounted a staunch opposition to the program as the 2022 general elections approached. FdI, the main opposition party and front-runner at the elections in particular, pledged to replace the measure with a program targeted at a smaller cohort of beneficiaries.

4.2.3 RdC and M5S

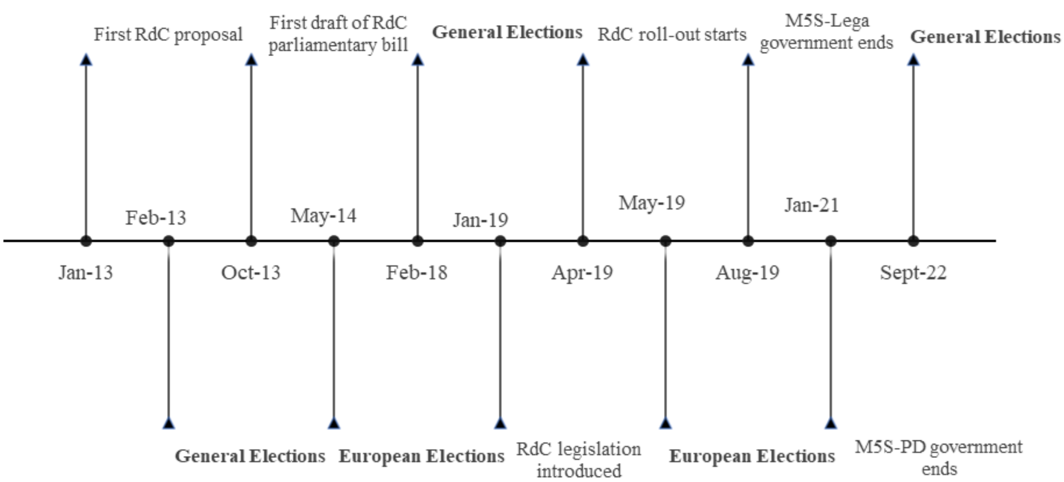
Officially founded in 2009 by the comedian Beppe Grillo, the M5S displays some of the typical features of a populist movement (Tronconi, 2015). Most importantly, it claims to represent "the people", assumed as unified by a common interest and juxtaposed with the corrupt establishment. At the local and regional elections of 2010, 2011, and 2012, the M5S obtained unexpected victories drawing from an electoral base predominantly formed by disappointed (center-)leftist voters (Pedrazzani and Pinto, 2017; Natale, 2014). In February 2013, almost nine million Italians chose the M5S, making it the most-voted party in the lower chamber of parliament. As mentioned above and reported in Figure 4.2.1, the RdC appeared for the first time as a policy proposal two weeks before the first M5S participation in a national election in 2013. The following years confirmed the M5S as a force to be reckoned with, with the election of seventeen MEPs at the European election in 2014 (21,5% of vote share) and the victories of its candidates as mayors in two of the country's largest cities (Rome and Turin) in 2016. It is however at the 2018 general election that the M5S obtains the highest level of electoral consensus to date with 32,8% of votes.
In the run-up to the 2018 election, the introduction of a guaranteed minimum income measure like the RdC represented a prominent policy proposal in the party electoral manifesto and one of the pillars in the communication strategy of its leaders. Analysis of media cov-

erage focusing on the two months ahead of the March 2018 elections highlights how the RdC proposal was the third most frequent economic policy proposal discussed in the Italian printed media among all measures put forward by the major parties competing in the elections (Cattaneo, 2018).

Having formed a coalition government with the far-right populist party Lega, in January 2019 the deputy prime minister and M5S party Leader, Luigi Di Maio, announced the introduction of the RdC. The program roll-out began in April of the same year, just two months ahead of the coming European elections, where the party registered a significant dip in consensus (17.1%).

In August of the same year, a vote of no confidence triggered by Lega MPs put an end to the first experiment of a populist-led government in Western Europe. In the following three years, two successive and distinct coalition governments were formed, both with M5S as senior partner. The former was led by the same prime minister and the center-leftist Democratic party replaced Northern League as a junior partner. The latter was instead supported by a broader parliamentary coalition and the new prime minister, Mario Draghi, had no party affiliation. None of the two governments questioned the RdC. Early elections were finally called in September 2022 after the M5S withdrew its support to Mario Draghi's government. As discussed in subsection 4.2.2, the run-up to the September 2022 general elections - where the M5S obtained 15.35% of the total share of votes - saw the RdC once again at the centre of the political debate, with the M5S pledging to expand the program in contrast to the promises to abolish it from the election front-runner and winner, FdI.

Figure 4.2.1: Timeline of main elections and RdC related events.



4.3 Conceptual framework

The conventional wisdom presumes that government transfers sway votes toward the incumbent party[2]. However, the mechanism behind such an effect is not always clear. Plausible explanations pin down different theories of voter behavior and are contingent on both the policy details and the politics of such transfers.

First, voters may support the incumbent when the continuation of the transfers depends on its continuation in power. This mechanism is more likely when the policy was not passed with broad support, when a single party claims credits for its implementation, and when a challenger party campaigns against the program, and this threat is perceived as credible. A second possible mechanism is based on a quid pro quo. This is absent when the policy is programmatic (Hicken, 2011), i.e. citizens receive transfers according to well-defined rules and not based on their partisanship or voting history. To be sure, even an exchange like vote-buying[3] is difficult to enforce because of the secrecy of the ballot. Third, voters may feel a sense of obligation to the party that favored them and cast their vote to reciprocate (Manacorda et al., 2011). Although some reciprocity is needed even in the case of clientelistic policies(Finan and Schetcher, 2012), it likely plays a more central role in explaining voters' behavior when transfers are programmatic, and therefore unconditional to voters' support. Reciprocity is also believed to be contingent on both the perceived value of the benefit received and the recipients' need (Gouldner, 1960). If this is the case both voters' expectations and their material welfare can make a difference.

For reciprocity to work, the nature of the politics under which a policy is passed and implemented also matters. Imai et al. (2020), distinguish between purely (or nonpartisan) and partisan programmatic redistributive policies. Both types of policies are programmatic as (i) their rules do not give incumbent parties discretion over implementation and (ii) voters receive the same services regardless of the party in office. Their differences stem from whether iii) they are passed with broad support from all political actors and whether iv) no single party claims exclusive credit for their implementation. In the review of two purely programmatic anti-poverty programs in Mexico, Imai et al. (2020) find no evidence of increasing electoral support for incumbent parties. Under nonpartisan programmatic conditions, indeed, voters believe that they would receive program benefits regardless of the incumbent's partisan identity. As evident from the discussion of its evolution as a policy proposal, on the contrary,

[2]See Golden and Min (2013) for a literature review

[3]Defined as a targeted attempt to weaken electoral discipline with relatively small individual transfer delivered personally (Finan and Schetcher, 2012).

the RdC fits the definition of a partisan programmatic policy as it did not receive bi-partisan support for its introduction, and a single party, the Five Stars Movement (M5S), claimed exclusive credit for its implementation. To this extent, an analysis of its impact on the electoral support for the M5S represents a unique case study for testing the "programmatic incumbent support hypothesis" when not all of the conditions of a purely programmatic policy are met.

Finally, to understand and analyse dynamics related to voters' disappointment, we draw on the literature on populist policy-making. As described by Gori (2020) and discussed in the previous sub-section, the procedural steps that led to the design and implementation of RdC present some of the key features of a populist policy. In line with Bartha et al. (2020)'s definition of populist policy-making, indeed, M5S officials involved in the design of the policy have i) downplayed the role of technocratic expertise, ii) have sidelined veto-players and iii) implemented fast and unpredictable policy changes, arguably with the aim of introducing the program ahead of the European elections. As supported by the evidence of this research, such features played a key role in explaining decisions made around the design and communication of the policy which in turn shaped voters' response to it.

4.4 Data and descriptive findings

Our empirical analysis is based on a panel data set of all Italian municipalities (n=7911) whose electoral results are observed in five nationwide elections. Three of them elected the lower chamber of national Parliament (2013, 2018 and 2022) and two representatives at the European Parliament (2014 and 2019). They represent so far the only nationwide elections with M5S participation. Our main outcome variable is the M5S share of valid votes.

To measure the intensity of RdC, we rely on information released by the National Social Security Institute (INPS) that provides both the number of beneficiaries and the monetary amount of benefits received at the municipality level in June 2019 and January 2020. We use the latter to extrapolate the same indicator for September 2022 based on variations in the share of RdC recipients observed at the province level. We then construct an indicator of treatment intensity, measured as the total number of RdC beneficiaries with Italian or EU nationalities[4] over Italian residents in each municipality. Thus, our indicator of treatment intensity is defined in each municipality m for $t = 2019$ and 2022

[4]We do not have information on the number of EU non-Italian citizens.

$$Recipients_{mt} = \frac{\#Recipients_{mt}}{\#Residents_{mt}}. \tag{4.1}$$

A first set of covariates relates to economic indicators made available by the Italian Ministry of Economics and Finance. They are based on annual individual tax returns aggregated at the municipality level. Most importantly, we use information that allows us to capture some of the eligibility criteria for RdC such as the share of taxpayers, the distribution of taxpayers into seven income-bands as well as the average estate income.

Finally, we use socio-demographic variables describing the number of foreign residents, the average household size, and the share of unemployed individuals in each municipality provided by the Italian National Institute of Statistics (ISTAT).

Table 4.4.1 provides the descriptive statistics for the treatment and control variables, each broken down between municipalities with treatment intensity below and above the median value. As expected, the incidence of RdC beneficiaries correlates with variables related to the eligibility requirements. In particular, municipalities with high RdC intensity show a higher share of households with no or low income, lower real estate income, and lower shares of foreign residents. Moreover, municipalities with a high incidence of RdC also show a higher rate of unemployment, have on average larger families and less educated residents.

The geographical distribution of municipalities with high and low RdC incidence is also uneven. As expected, in a country characterised by economic dualism, a high share (87.8%) of municipalities with a low RdC incidence is located in the North, whereas 60.3% of those with a high incidence are located in the South.

Table 4.4.2 presents the descriptive statistics for our main outcome variable. It provides a breakdown by each of the four elections considered in the study, further divided among municipalities with low and high RdC intensity. This breakdown allows us to observe how, in 2013, the M5S share of votes was evenly distributed across municipalities with high and low RdC intensity (and therefore other socio-economic characteristics discussed above), with an average support of only 1.9pp higher in municipalities with an RdC incidence above the median value. This gap more than doubled to 3.9pp at the 2014 European elections, to reach 13.3pp at the 2018 general elections. In 2019, the party almost halved its electoral consensus and while the gap between the two groups of municipalities remained substantial, it dropped to 11pp, to rise again slightly at the 2022 general elections to 11.3pp.

Table 4.4.1: Descriptive statistics of independent variables

	Mean	SD	Min	Max
RdC intensity	0.027	0.026	0.000	0.212
Low RdC	*0.008*	*0.005*	*0.000*	*0.017*
High RdC	*0.045*	*0.026*	*0.017*	*0.212*
Eligibility Controls				
No income	0.162	0.079	0.000	0.508
Low RdC	*0.114*	*0.051*	*0.000*	*0.430*
High RdC	*0.209*	*0.075*	*0.000*	*0.508*
Income 0-10k	0.259	0.075	0.096	0.952
Low RdC	*0.228*	*0.064*	*0.130*	*0.952*
High RdC	*0.291*	*0.071*	*0.096*	*0.628*
With estate income	0.371	0.084	0.050	0.625
Low RdC	*0.413*	*0.069*	*0.088*	*0.625*
High RdC	*0.330*	*0.076*	*0.050*	*0.556*
Estate income	1.024	640	77	28.209
Low RdC	*1.109*	*745*	*77*	*28.209*
High RdC	*940*	*502*	*84*	*12.709*
Foreigners	0.068	0.044	0.000	0.417
Low RdC	*0.074*	*0.040*	*0.000*	*0.417*
High RdC	*0.062*	*0.046*	*0.000*	*0.336*
Other controls				
Unemployment	0.120	0.062	0.009	0.394
Low RdC	*0.080*	*0.030*	*0.009*	*0.322*
High RdC	*0.160*	*0.060*	*0.011*	*0.394*
Avg. family size	2.301	0.255	1.115	3.415
Low RdC	*2.291*	*0.238*	*1.115*	*3.056*
High RdC	*2.312*	*0.269*	*1.167*	*3.415*
Uni degree	0.078	0.029	0.000	0.289
Low RdC	*0.076*	*0.029*	*0.000*	*0.289*
High RdC	*0.079*	*0.029*	*0.009*	*0.260*
North	0.555	0.497	0.000	1.000
Low RdC	*0.878*	*0.328*	*0.000*	*1.000*
High RdC	*0.235*	*0.424*	*0.000*	*1.000*
Centre	0.123	0.328	0.000	1.000
Low RdC	*0.084*	*0.277*	*0.000*	*1.000*
High RdC	*0.161*	*0.368*	*0.000*	*1.000*
South	0.322	0.467	0.000	1.000
Low RdC	*0.038*	*0.192*	*0.000*	*1.000*
High RdC	*0.603*	*0.489*	*0.000*	*1.000*

Notes: Number of observations 9,711. *RdC intensity* is the share of RdC beneficiaries in each municipality. *Low RdC* are the municipalities with RdC share below the median value. *High RdC* are municipalities with RdC intensity above its median value. *No income* is the share of over-18 residents reporting no taxable income. *Income 0-10k* is the share of residents with yearly taxable income below €10,000. *With estate income* is the share of residents reporting income from real estate. *Estate income* is the average real estate income expressed in euros. *Foreigners* is the share of foreign residents. *Unemployment* is the share of unemployed residents. *Avg family size* indicates the average family size in each municipality. *Uni degree* is the share of residents with a university degree, *North, Centre, South* indicate the macro area of each municipality.

Table 4.4.2: Descriptive statistics of dependent variable

	Mean	SD	Min	Max
M5S Vote Share				
2013 General elections	0.237	0.071	0.001	0.581
Low RdC	*0.228*	*0.070*	*0.001*	*0.581*
High RdC	*0.247*	*0.070*	*0.044*	*0.563*
2014 EU elections	0.195	0.065	0.000	0.633
Low RdC	*0.176*	*0.057*	*0.000*	*0.494*
High RdC	*0.215*	*0.067*	*0.018*	*0.633*
2018 General elections	0.297	0.117	0.000	0.736
Low RdC	*0.230*	*0.075*	*0.000*	*0.711*
High RdC	*0.363*	*0.113*	*0.000*	*0.736*
2019 EU elections	0.157	0.093	0.000	0.624
Low RdC	*0.102*	*0.051*	*0.000*	*0.498*
High RdC	*0.212*	*0.092*	*0.000*	*0.624*
2022 General elections	0.130	0.095	0.000	0.851
Low RdC	*0.073*	*0.042*	*0.000*	*0.851*
High RdC	*0.186*	*0.099*	*0.000*	*0.593*

Notes: Number of observations 9,711. *M5S vote share* is the share of valid votes obtained by the M5S movement in each municipality at the 2013, 2018 and 2022 elections for the lower house of the Italian parliament (*2013 General elections*, *2018 General elections* and *2022 General elections* respectively) and at the 2014 and 2019 European Parliament elections (*2014 EU elections* and *2019 EU elections* respectively). *Low RdC* are the municipalities with RdC share below the median value. *High RdC* are municipalities with RdC intensity above its median value.

A clear trend emerges from this descriptive overview: support for the M5S is higher in areas where the incidence of RdC recipients is higher. Indeed, by pooling together election data for the years following the introduction of the policy (2019 and 2022) in a simple OLS model we find a positive and statistically significant correlation between the share of beneficiaries and the M5S electoral performance at the municipality level. As illustrated in Table 4.4.3, a 1pp increase in the share of beneficiaries is associated with an increase in the share of M5S votes ranging from 1.94pp to 0.68pp once we control for economic variables related to the eligibility criteria of the measure and include a province level dummy. The descriptive evidence presented in Table 4.4.2, however, indicates how this trend seems to predate the introduction of the RdC, with the party consistently increasing its support base in relatively poorer areas. For this reason, we can not consider the correlation observed here as the effect of introducing the RdC on the M5S electoral outcomes, and a more sophisticated strategy is needed.

Table 4.4.3: Panel regression, European elections 2019 and general elections 2022

VARIABLES	(1) M5S votes share	(2) S votes share	(3) 5S votes share
% Beneficiaries	1.941***	1.273***	0.681***
	(0.0206)	(0.0259)	(0.0237)
Year = 2019	0.106***		
	(0.000929)		
Year = 2022	0.0667***	-0.0320***	-0.0317***
	(0.00100)	(0.000531)	(0.000518)
Constant		-0.0660***	0.240***
		(0.00712)	(0.0106)
Observations	15,608	15,601	15,601
Number of codicecomune	7,900	7,899	7,899
Controls	No	Yes	Yes
Province dummies	No	No	Yes

Notes: Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. The dependent variable is the share of M5S vote share at the 2019, and 2022. *%Beneficiaries* is the share of RdC recipients in the municipality over the voting eligible population. Columns (2) and (3) include the following controls: % of adult population with no taxable income the % of adult population with income less than €10,000 per year, % of population with estate income and the average estate income.

4.5 Empirical framework

In our analysis, we use municipality level data that combine the intensity of the RdC and the actual political support received by M5S. Our basic estimation strategy relies on a standard two-way fixed effects regression:

$$Y_{mt} = \theta_t + \eta_m + \beta \cdot Recipients_m \cdot Post_t + v_{mt}, \tag{4.2}$$

where Y_{mt} denotes the percentage of votes cast for M5S over total votes in municipality m in election years t =2018, 2019 and 2022, θ_t and η_m represent respectively time and municipality fixed effects, $Recepients_{mt}$ is our measure of treatment intensity defined in equation 4.1, and $Post_t$ is a dummy for the post treatment period. Our coefficient of interest is β and captures the relationship between RdC intensity and M5S vote share. The standard errors of our regressions are clustered at the municipality level.

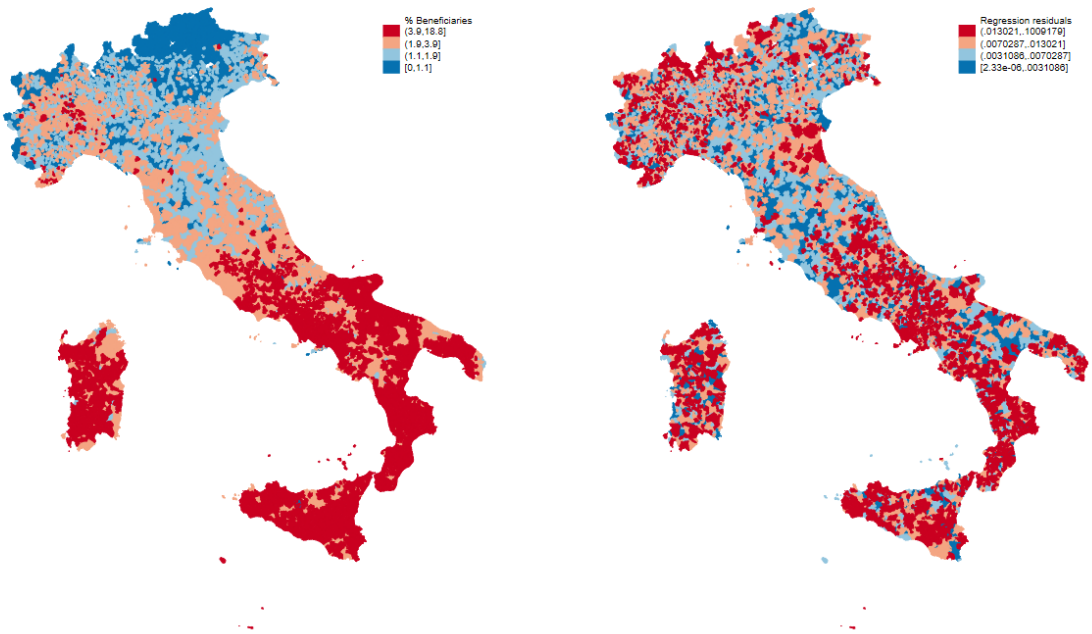
Identifying the electoral impact of anti-poverty programs is challenging. Even with the appropriate individual data that combine transfer receipt with voting behaviour, omitted variable, and reverse causality are likely: first, social groups that benefit the most from these programs have often ex-ante different political preferences (e.g. for parties that favor redistribution). Second, politicians are likely to target specific groups such as core supporters or swing voters. Hence a simple correlation between the RdC intensity and M5S electoral results is not always informative on the causal effect of the policy on political support. To address these problems, we apply a difference-in-differences design. More specifically, we use equation 4.2 to compare the changes over time in electoral results of M5S in municipalities with different intensities of RdC. Admittedly, even if this approach allows to control for time-invariant municipality characteristics, it faces a further set of challenges.

First, differently from the "canonical" difference-in-difference model, our treatment is continuous, and hence to interpret $beta$ (equation 4.2) as a causal response parameter, we need a "stronger" parallel trends assumption. In particular, we need to assume that, for all intensities (or doses) of treatment, the average change in outcomes over time across all units if they had been assigned that amount of dose is the same as the average change in outcomes over time for all units that experienced that dose (Callaway et al., 2021). Note that, differently from the standard parallel trend assumption that concerns only untreated potential outcomes, this one involves potential outcomes under different doses.

A second challenge to the parallel trend assumption concerns the possibility that, given the heterogeneity in dimensions extending beyond the distribution of RdC across municipalities with different treatment intensities, contemporary shocks may affect differently municipality with more versus fewer recipients. For example, as depicted in Table 4.4.1, municipalities with more recipients are predominantly located in the South. Hence, they may differ in dimensions that affect the change of M5S electoral outcomes over time, leading to an infringement of the standard parallel trend assumption. To alleviate this concern, we exploit the eligibility requirements of RdC. As mentioned in Section 4.2.1, to be eligible households must meet four criteria. The most important ones are (i) having an ISEE index lower than

€9,360 and (ii) not owning a real estate property whose value exceeds €30,000. We can control separately for municipality-level variables that capture the two requirements while still capturing variation in the actual intensity of RdC. Doing so allows us to compare municipalities with different exposure to the treatment as the result of the interaction of two or more eligibility criteria, but otherwise similar along the other dimensions considered in the model, thus capturing quasi-exogenous variations in treatment intensity (Vannutelli, 2023). As mentioned in section 4.4, data from tax records provide us with information that can be used as proxy measures for both income and real estate indicators at the municipality level, allowing us to control for the main eligibility criteria in our specification model. A graphical demonstration of the quasi-exogenous variation in treatment intensity exploited in our identification strategy is displayed in Figure 4.5.1. Panel A plots the distribution in the share of RdC beneficiaries at the municipality level, while Panel B shows the residual variation, after controlling for the income distribution and the real estate indicators. While a clear geographic pattern is present in the former, this is less visible on the right, suggesting the importance of including the eligibility-related controls.

Figure 4.5.1: Distribution of RdC beneficiaries at the municipality (Panel A) and of the residual variation, after controlling for the eligibility threshold (Panel B)



Building on equation 4.2, our second specification includes a set of eligibility criteria controls as follows:

$$Y_{mt} = \theta_t + \eta_m + \beta \cdot Recipients_{mt} \cdot Post_t + \gamma \cdot X_{mt} \cdot Post_t + v_{mt}, \qquad (4.3)$$

where X_{mt} is a matrix of controls capturing the different eligibility criteria (i.e the % of non-tax payers, the % of individuals with yearly income below €10,000, % of individuals with real estate income, average real estate income) interacting with the dummy post which is equal to one in the elections years 2019 and 2022, following the introduction of the RdC.
Finally, to test the robustness of our results, we exploit pre-treatment periods to relax the common trend assumption (Angrist and Pischke, 2014). As shown in the previous Section, pre-treatment outcomes trends have not been parallel between municipalities with different doses of treatment: from year to year, M5S electoral electoral results improved in municipalities with high RdC intensity relative to municipalities with low RdC intensity. Note that this was the case also in the 2014 European election when the overall M5S result was lower than the 2013 National one. Nonparallel trends before the treatment are suggestive of nonparallel trends after the treatment (in the absence of the treatment). For this reason, we add to equation 4.3 a term for municipality-specific trends $m \cdot t$ as described in equation 4.4 presuming that, in the absence of the treatment, the M5S results would have deviated from common year effects by following a linear trend within each municipality. In addition, we expand t to include all nationwide election years in which the M5S took part, t=2013 and t=2014:

$$Y_{mt} = \theta_t + \eta_m + \beta \cdot Recipients_{mt} \cdot Post_t + \gamma \cdot X_{mt} \cdot Post_t + \lambda \cdot m \cdot t + v_{mt}. \qquad (4.4)$$

.

4.6 Results

We first apply our difference-in-differences strategy to the full sample of elections following the introduction of the RdC. If we consider the most basic model (Column 1 in Table 4.6.1, for each percentage point increase in the share of beneficiaries across municipalities, the share of M5S votes drops by 0.137pp. However, once the eligibility criteria are included (Column 2), the effect of the introduction of the RdC becomes positive, leading to a 0.2pp increase in the support for the M5S for every percentage point increase in the share of RdC beneficiaries. If we relax the parallel trend assumption by exploiting electoral data from all elections the

M5S partook in and by controlling for non-parallel trends at the municipality level (Column 3), we find that one percentage point increase in the local share of RdC beneficiaries is linked to a 1.14pp drop in the share of votes for the M5S. Based on these findings, we cannot establish whether the introduction of the RdC did in fact have an impact on the electoral support of the M5S.

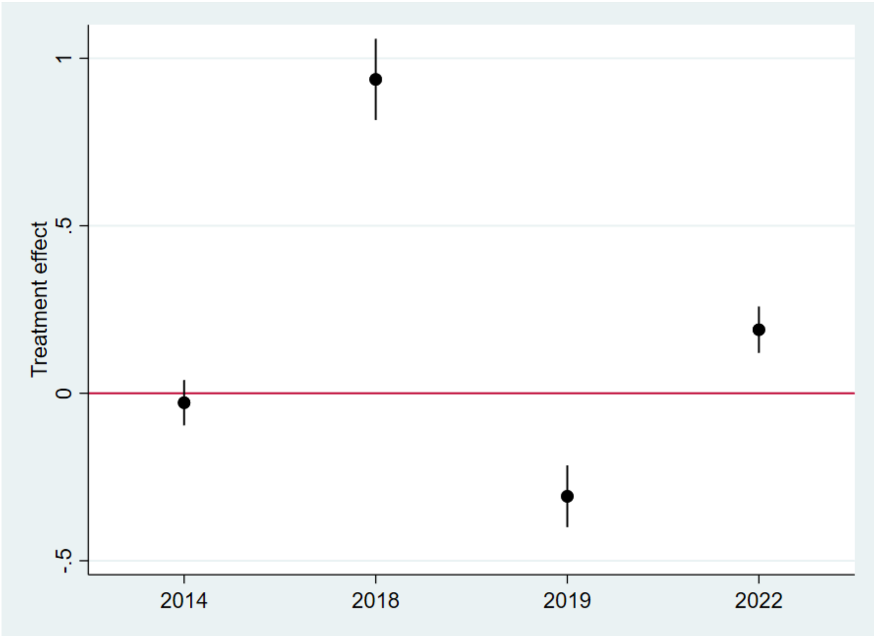
Table 4.6.1: Diff-in-Diff results, general elections 2022

VARIABLES	(1) % voti M5S	(2) % voti M5S	(3) % voti M5S
% Beneficiaries	-0.137***	0.191***	-1.143***
	(0.0238)	(0.0354)	(0.0332)
Year = 2014			-0.0467
			(0.624)
Year = 2018			0.0358
			(3.121)
Year = 2019	-0.136***	-0.0409***	0.00916
	(0.000773)	(0.00842)	(3.745)
Year = 2022	-0.164***	-0.0728***	-0.0424
	(0.000850)	(0.00835)	(5.618)
Constant	0.298***	0.298***	0.258
	(0.000370)	(0.000360)	(2.607)
Observations	23,508	23,500	39,301
R-squared	0.945	0.948	0.930
Muncipaity Clustering	Yes	Yes	Yes
Controls	No	Yes	Yes
Muncipality-level trends	No	No	Yes

$Notes$: Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. The dependent variable is the share of M5S vote share at the 2018, 2019 and 2022 elections in Columns (1) to (3), and at the 2013, 2014, 2018, 2019 and 2022 elections in Column (4). $\%Beneficiaries$ is the share of RdC recipients in the municipality over the voting eligible population. Columns (2) to (4) include the following controls: % of adult population with no taxable income the % of adult population with income less than €10,000 per year, % of population with estate income and the average estate income.

A key potential limitation in our identification strategy relates to the possible presence of anticipatory effects. As previously discussed in Section 4.4, the introduction of RdC was largely anticipated and the 2018 elections determined a crucial shift in the distribution of the M5S electoral support which saw their consensus concentrating in the most economically deprived areas of the country. To further explore this possibility, we carry out a placebo test in a difference-in-difference setting for both the 2014 and 2018 elections. Figure 4.6.1 plots the regression coefficients of the share of beneficiaries on the difference in M5S electoral support in the 2014, 2018, 2019, and 2022 elections, controlling for the eligibility criteria as described in equation 4.3. While the coefficient is not statistically significant for the 2014 European elections, we find a strong positive and statistically significant coefficient for the 2018 general elections. This finding suggests that the assumption of no anticipatory effects is in fact problematic. It's important to remember what was discussed in Section 4.2 in relation to how the electoral campaign leading up to the February 2018 vote was heavily centered around the RdC proposal which likely swayed many voters, who self-identified as potential beneficiaries, to vote for the M5S. Interestingly, and perhaps relatedly, the following elections (the first after the implementation of the RdC) present an opposite relation between the M5S performance and the incidence of the RdC, while reverting back to positive and statistically significant at the 2022 general elections. The combined effects of these contrasting trends are likely the reason behind the inconsistent findings presented above. In light of these results and given the complex nature of the politics which accompanied the RdC at every stage of its evolution, from policy proposal to highly contested enacted measure, we set out to examine the dynamics behind each election separately.

Figure 4.6.1: Regression coefficients of RdC incidence on M5S electoral outcomes at the 2014 and 2019 European elections and 2018 and 2022 general elections.



4.6.1 Mechanism - Disappointment at the European 2019 elections

The 2019 European elections, held in May, were the first electoral contest held after the introduction of the RdC in April of that year. Gori (2020) described the timeline of events as strictly intentional, with the incumbent party rushing to implement the RdC into policy with the hope of capitalising on consensus for enacting their flagship electoral proposal. For this reason, assessing the introduction of the RdC on the M5S electoral outcome at the 2019 European elections is relevant to understand the extent to which the party's strategy was successful.

By applying the same difference-in-difference approach of equation 4.3 to include only the 2018 and 2019 elections, our findings show a negative correlation in the share of RdC recipients on the M5S vote share at the 2019 European elections. Table 4.6.2 (Column 2) shows how this coefficient persists with the introduction of the eligibility threshold controls. A 1 pp. increase in the share of RdC beneficiaries is linked to a 0.22 pp. drop in support for the M5S between the 2018 and 2019 elections. In June 2019, the average share of RdC recipients among Italian residents was of 2.6%. Approximating by this figure, we can conclude that the introduction of the RdC is associated, on average to a 0.59 pp. drop in the M5S votes share between the 2018 and 2019 elections. The results are robust to a set of further checks.

First, we can account for other exogenous shocks that occurred between the two elections, such as the decision of the M5S to form a coalition government with the League, by including an area-level dummy fixed effect in the model, (Column 3). Second, we can relax the assumption of parallel trends by including all elections dating back to 2013 and inserting municipality-level yearly trends (Column 4). Third, in order to control for any difference in voters' behaviour between the general and European elections, we compare the difference between European and general elections in 2019 and 2018 with those of 2014 and 2013, by estimating a triple difference as our dependent variable (Column 5). The presence of a negative and statistically significant coefficient persists across all the above specifications.

Table 4.6.2: Diff-in-Diff results 2019 European elections

VARIABLES	(1) % M5S votes	(2) % M5S votes	(3) % M5S votes	(4) % M5S votes	(5) Diff-Diff-Diff
% Beneficiaries	-0.454***	-0.226***	-0.0875**	-0.396***	-0.202***
	(0.0256)	(0.0425)	(0.0408)	(0.0368)	(0.0567)
Year = 2019	-0.127***				
	(0.000725)				
Constant	0.297***	0.297***	0.297***	0.464	-0.0536***
	(0.000263)	(0.000257)	(0.000250)	(4.421)	(0.00900)
Observations	15,800	15,792	15,792	31,596	7,896
R-squared	0.966	0.967	0.969	0.951	0.168
Municipality Clustering	Yes	Yes	Yes	Yes	No
Controls	No	Yes	Yes	Yes	Yes
Area fixed-effects	No	No	Yes	No	No
Municipality-level trends	No	No	No	Yes	No

Notes: Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. The dependent variable is the difference in M5S vote share between 2019 and 2018 elections for columns (1)-(4) and the double difference in M5S vote shares between 2019 and 2018 elections and 2014 and 2013 elections in column (5). % Beneficiaries is the share of RdC recipients in the municipality in June 2019 over the voting eligible population. Columns (2) to (5) include the following controls: % of adult population with no taxable income the % of adult population with income less than €10,000 per year, % of population with estate income and the average estate income.

The presence of an anticipation effect at the 2018 general elections leads us to consider how the negative correlation observed in 2019 is reflective of widespread disappointment among those voters who received the benefit. Aspects related to the policy design and its populist nature are likely to have generated expectations that were successively unmet once the measure was introduced. As previously discussed, the introduction of the RdC was preceded by an electoral campaign in which the measure featured as one of the most prominent policy proposals. Party officials made frequent reference to the monthly sum of €780 as the actual amount of cash handout for eligible households should the policy be implemented. This messaging was a legacy of the policy proposal presented to the parliamentary chambers in October 2013, which established the sum of €780 as the monthly income threshold for households to be eligible. Bound by these electoral promises, M5S officials elected in government ensured that the sum was still a feature of the implemented policy by establishing it as the maximum amount available to single-unit households with an ISEE score of zero. In order to fulfill this promise while maintaining the overall expenditure of the program within the limits agreed with their coalition partners, the trade-off was obtained at the expense of larger families with the introduction of a relatively less generous equivalence scale. Further to this, the fixed amount of rental support made available is too a feature of the policy set to disadvantage more numerous families who live in larger properties and therefore face higher rents. The result of these choices was that two months from the official rollout of the scheme, the average benefit amount awarded was €450 per month for the 1m beneficiaries across the country, well below the €780 figure so prominently present in the movement communication campaign. Based on these considerations we expect large numbers of beneficiaries to be left disappointed by the amount of benefits received. We set out to test this hypothesis in two ways. Firstly we include in our set of difference-in-difference specifications a variable describing the average amount of RdC per recipient in each municipality. By doing so we can test whether, by holding the share of beneficiaries constant, support for the M5s varies in a significant way in line with variations in the average amount of benefits awarded to each individual. Table 4.6.3 reports the findings of the regressions. The positive coefficient of the RdC amount variable in all but one column included seems to support our hypothesis: holding the share of recipients in each municipality constant, the support for the M5S increases as the average amount of benefit awarded increases. The analysis thus suggests that the coefficient of the treatment variable, the share of beneficiaries, is related to the average amount awarded and therefore the negative overall effect observed on the electoral support for the M5S at the 2019 European elections can be partly explained by the lower-than-expected amount of benefits awarded.

Table 4.6.3: Diff-in-Diff results with average RdC amount

VARIABLES	(1) % M5S votes	(2) % M5S votes	(3) % M5S votes	(4) % M5S votes	(5) Diff-Diff-Diff
% Beneficiaries	-0.453***	-0.222***	-0.0836**	-0.390***	-0.194***
	(0.0255)	(0.0426)	(0.0408)	(0.0368)	(0.0567)
RdC amount	-5.27e-06**	4.16e-06	4.39e-06*	6.47e-06**	8.73e-06**
	(2.56e-06)	(2.55e-06)	(2.43e-06)	(2.64e-06)	(3.48e-06)
Year = 2019	-0.125***				
	(0.00146)				
Constant	0.297***	0.297***	0.297***	0.464	-0.0543***
	(0.000263)	(0.000257)	(0.000250)	(4.420)	(0.00901)
Observations	15,800	15,792	15,792	31,596	7,896
R-squared	0.966	0.967	0.969	0.951	0.169
Muncipaity Clustering	Yes	Yes	Yes	Yes	No
Controls	No	Yes	Yes	Yes	Yes
Muncipality-level trends	No	No	No	Yes	No
Area fixed-effects	No	No	No	Yes	No

Notes: Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. The dependent variable is the difference in M5S vote share between 2019 and 2018 elections for columns (1)-(4) and the double difference in M5S vote shares between 2019 and 2018 elections and 2014 and 2013 elections in column (5). % Beneficiaries is the share of RdC recipients in the municipality in June 2019 over the voting eligible population. $RdCamount$ is average RdC monthly amount awarded per recipient in municipality m. Columns (2) to (5) include the following controls: % of adult population with no taxable income the % of adult population with income less than €10,000 per year, % of foreign-born, % of population with estate income and the average estate income.

In a similar fashion, we explore this mechanism further by focusing on the impact of specific features of the policy design on the electoral behaviour of beneficiaries in households of varying sizes. By adopting an equivalence scale biased in favour of single households and by establishing a flat amount of rental support, policymakers effectively introduced a measure that leaves larger households with monetary support below their economic needs. As a result of this, we expect beneficiaries voters living in larger families to show relatively less support for the M5S than beneficiaries voters living in smaller households. We test this hypothesis by inserting an interaction term between the share of beneficiaries and the average household size of each municipality. Table 4.6.4 reports a negative and statistically significant coefficient in the interaction terms which supports our hypothesis further. Holding the share of beneficiaries as constant, electoral support for the M5S at the 2019 European elections has decreased in municipalities with larger average household sizes where voters expressed their disappointment with regard to the insufficient amount of benefits received by withdrawing their support from the incumbent party. These findings offer clear empirical evidence of how, despite investing an unprecedented amount of public funds in a redistributive policy aimed at tackling poverty, poor choices around the communication strategy and the policy design resulted in widespread disappointment among RdC beneficiaries reflected in a loss of electoral support for the M5S. Moreover, they offer an example of how the principle of reciprocity is indeed contingent on both the perceived value of the benefit received and the recipients' needs, which in this case were largely perceived as unmet.

Table 4.6.4: Diff-in-Diff results. Interaction with average family size

VARIABLES	(1) % M5S votes	(2) % M5S votes	(3) % M5S votes	(4) % M5S votes	(5) Diff-Diff-Diff
% Beneficiaries	0.0822	0.330*	0.531***	0.580***	0.468***
	(0.176)	(0.173)	(0.163)	(0.135)	(0.159)
% Benefic. * Family Size	-0.216***	-0.228***	-0.255***	-0.397***	-0.269***
	(0.0715)	(0.0712)	(0.0665)	(0.0535)	(0.0631)
Year = 2014				-0.118	
				(1.468)	
Year = 2018				-0.319	
				(7.339)	
Year = 2019	-0.128***	-0.0516***		-0.436	
	(0.000734)	(0.00781)		(8.806)	
Constant	0.297***	0.297***	0.297***	0.464	-0.0734***
	(0.000263)	(0.000257)	(0.000249)	(4.403)	(0.00820)
Observations	15,798	15,792	15,792	31,596	7,896
R-squared	0.966	0.967	0.969	0.951	0.181
Muncipaity Clustering	Yes	Yes	Yes	Yes	Yes
Municipality-level trends	No	No	No	Yes	No
Area fixed-effects	No	No	Yes	No	No

$Notes$: Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. The dependent variable is the difference in M5S vote share between 2019 and 2018 elections for columns (1)-(4) and the double difference in M5S vote shares between 2019 and 2018 elections and 2014 and 2013 elections in column (5). $\%Beneficiaries$ is the share of RdC recipients in the municipality in June 2019 over the voting eligible population. $\%Benefic.*FamilySize$ is the interaction term between the share of RdC recipients and the average family size in each municipality. Columns (2) to (5) include the following controls: % of adult population with no taxable income the % of adult population with income less than €10,000 per year, % of foreign-born, % of population with estate income and the average estate income.

4.6.2 Mechanisms - In defense of the policy at the 2022 general elections

Contrary to what was observed in 2019, our main findings suggest that the impact of the RdC on the M5S electoral outcome at the 2022 general elections is positive. Our hypothesis is that differences in the observed results between the two elections stem from differences in the nature and the stakes of the elections themselves. While the former, held right after the rollout of the RdC, did not yield direct consequences in relation to domestic issues, the latter represented an electoral contest in which the future of the RdC was undoubtedly at stake. As discussed in Subsection 4.2.2, the lead-up to the vote was characterised by a vocal campaign by the elections front-runner and eventual winner, the right-wing party *Brothers of Italy* (FdI), to replace the RdC with a new program that would reduce the number of beneficiaries to a smaller cohort. Amidst this polarizing debate around the future of the measure, our hypothesis is that voters in receipt of the RdC, faced with the concrete threat of seeing their benefits withdrawn, turned once more towards the main political force in support of the measure, the M5S.

To test this, we exploit the contemporary unfolding of two distinct elections held in the region of Sicily. In September 2022, Sicilian voters were called to elect, alongside their representatives in the national parliament, the president of the regional government as well as the representatives of the regional assembly. All main parties, including the M5S, featured on both ballots and the overall alignment across the different coalitions at the regional level reflected the national political picture. Given the similarity in the options available and the contemporaneity of the voting, we believe that differences in the outcomes between the two elections stem from voters' perceptions of parties' different stands with respect to national and regional issues, the RdC being one of them. A Sicilian voter in receipt of the RdC was thus confronted with two ballots to be cast at the same time, both presenting almost identical voting options. The main difference, among others, is that on the general elections ballot the continuation of the RdC was at stake.

Hence, we use regional election outcomes as a counterfactual and check whether the difference in the electoral outcomes of the M5S between the national and regional elections is correlated with the distribution of RdC beneficiaries across Sicilian municipalities. Table 4.6.5 reports the results of these regressions. The most conservative estimates (Column 3) indicate how for each percentage point increase in the incidence of RdC beneficiaries the vote share of the M5S increased by 0.2pp between the regional and national elections. Given the party overall difference of 13pp between the two contests (28.2% vs 15.2% at the general and regional elections respectively) and the mean incidence of RdC beneficiaries in Sicily of 12%, we estimate that the RdC accounts for between 6.18 pp and 2.5 pp in the difference of M5S share of votes. While these findings are related to one of Italy's 20 regions, the similarity in the magnitude of the coefficients reported in Table 4.6.5 (Columns 2 and 3) and Table 4.6.1 (Column 2) is consistent with the results obtained at the national level.

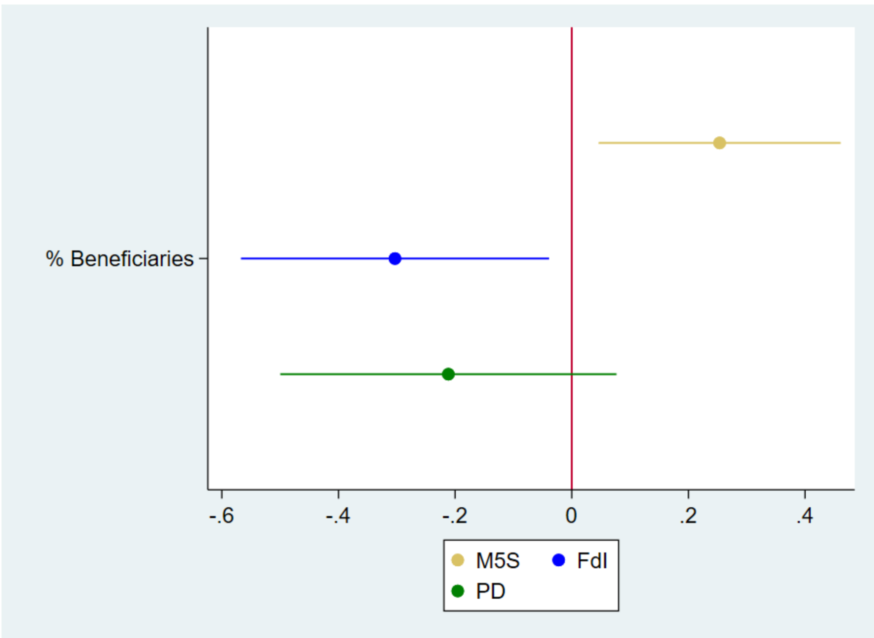
Table 4.6.5: Regression results 2022 general elections and regional elections in Sicily

VARIABLES	(1) $\Delta\%M5SSicily$	(2) $\Delta\%M5SSicily$	(3) $\Delta\%M5SSicily$
% Beneficiaries	0.515***	0.254**	0.204**
	(0.0727)	(0.106)	(0.0993)
% No income		0.324***	-0.0190
		(0.0753)	(0.0822)
% Pop. 0-10k		0.0204	-0.122*
		(0.0708)	(0.0706)
% Pop. estate income		0.112*	-0.100
		(0.0594)	(0.0628)
Average estate income		-3.22e-05***	-3.02e-05***
		(7.60e-06)	(7.21e-06)
Constant	0.0915***	0.0168	0.255***
	(0.00708)	(0.0475)	(0.0536)
Observations	390	390	390
R-squared	0.114	0.192	0.348
Controls	No	Yes	Yes
Province fixed effects	No	No	Yes

Notes: Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. The dependent variable is the difference in the M5S vote share between the 2022 general and regional elections in Sicily . $\%Beneficiaries$ is the share of RdC recipients in the municipality over the voting eligible population. Columns (3) to (4) include the following controls: % of adult population with no taxable income the % of adult population with income less than €10,000 per year, % of population with estate income and the average estate income. Column (4) includes a dummy variable for each of Sicily's nine provinces.

In addition, we carry out an analysis of the flow of votes between the two elections held in Sicily in September 2022 to explore whether we observe related trends for parties with different stands and pledges in relation to the RdC. Figure 4.6.2 plots the coefficients of the treatment variable share of RdC beneficiaries regressed on the variation in the share of votes between the 2022 general and regional elections in Sicily for Italy's three largest parties, namely Brothers of Italy (FdI), the Democratic Party (PD) and M5S. For PD, a party that voted against the introduction of the policy in 2019 but later supported the program, we find no effect. On the contrary, our analysis suggests how for FdI, which nationally obtained over 26.0% of votes and ran on a manifesto pledging to replace the RdC with a less generous program, the difference in electoral outcomes between the two elections is negatively correlated with the local incidence of RdC beneficiaries. The similar magnitude in the coefficient of the opposite sign compared to what was observed for the M5S indicates a trend among Sicilian voters in receipt of the RdC away from FdI on the national ballot, supporting our hypothesis of a "tactical" and self-interested voting behaviour by RdC recipients once the cash transfer program is at stake.

Figure 4.6.2: Regression coefficients of RdC intensity on the difference between the electoral outcomes of the three main parties at the 2022 general and regional elections across 390 municipalities in Sicily.



4.7 Conclusion

This chapter provides an account of the political consequences of the recent process of polarization which characterised the public debate around poverty and the tools to contrast it in Italy over the course of the last decade. In particular, it described the process leading the M5S, an anti-establishment and populist party, to become the first political force among voters in the most economically disadvantaged areas of the country. In trying to assess the causal impact of the RdC on the M5S electoral outcome, it unveiled how voters responded to the electoral promise of a generous cash transfer program by supporting the party at the 2018 general elections and successively expressed their disappointment with how the program was enacted by withdrawing their support at the 2019 European elections. Our empirical strategy suggests how such disappointment is systematically correlated with the average family size at the municipality level, as the result of a communication strategy and policy design biased towards single households. Under the pledge to replace the RdC with a less generous program made by the main opposition party, however, we find a positive effect of the RdC on the M5S electoral outcome at the 2022 general elections. The findings of this study contribute important evidence to the vast literature on redistributive policies and voting behavior. First, they support the argument that the politics behind these measures play a crucial role (Imai et al., 2020). In the case of a partisan programmatic policy like the RdC, voters seem to clearly respond to parties' stands on the policy and cast their preferences accordingly. Whether this is reflected in support or not for the incumbent, however, it depends on the interacting combination of the type of elections (i.e. whether the policy is at stake on the ballot) and the extent to which voters' expectations were met. On the latter, the findings emerging from this case study indicate how aspects associated with populist policy-making can in fact backfire when voters perceive that electoral pledges are unmet. Overall these findings suggest how empirical studies on the political consequences of redistributive policies cannot do without an analysis of the characteristics and nature of such policies, as well as considerations of the characteristics of the very elections.

Chapter 5

Conclusions

This doctoral thesis studies the evolving landscape of poverty in Italy and assesses the targeting performance and political consequences of the largest national anti-poverty program, the RdC. Across the different chapters, we investigate how poverty is produced and reproduced by geographical inequalities, shedding light on the nexus between de-industrialisation, in-work and absolute poverty as well as social welfare schemes. We then highlight the political economy of such schemes by investigating their political consequences in the context of increasing polarisation around the issue of poverty and welfare schemes. Throughout the chapters, we apply statistical and econometrics models to survey and administrative data to provide evidence at a granular geographical level, thus expanding the contribution of data science to social science research.

Chapter 2 acknowledges how the phenomenon of in-work poverty and the process of de-industrialisation are interrelated. In particular, we provide a fine spatial mapping of how in-work poverty evolved across non-administrative small-area boundaries and correlate its distribution by changes in local sectoral employment in the context of de-industrialisation. Chapter 3 focuses on the Citizen Income (RdC) as the most extensive anti-poverty program ever introduced in the country. Using small-area estimation techniques, it provides novel and policy-relevant insights into the targeting performance of the scheme after its introduction in 2019. Moreover, the analysis provides tangible suggestions on how the policy could be improved to ensure it reaches a wider cohort of households living in poverty. Chapter 4 looks at the political consequences of introducing the RdC. We study how the partisan nature of the measure shaped voters' responses to it, with a focus on the electoral performance of the 5 Star Movement, the party which introduced it. We find that, in a context of high political polarization around this anti-poverty measure, voters are indeed sensitive to electoral pledges, form expectations accordingly, and cast their votes strategically. Overall, the present analyses depict the state-of-the-art of Italy's efforts to reduce poverty and raise questions about its dimensions of policy and politics.

Several policy implications emerge from these results. First, low-pay work and in-work poverty are linked to both the local demand for labour and the type of jobs. Employment programs and active labour market policy on one hand, and industrial policies on the other must consider the wage implications alongside overall employment effects as part of their interventions. Second, anti-poverty cash transfers should take into account the geographical heterogeneity of needs, for example by linking housing support to regional rent and property prices and removing local barriers to accessing support. Third, during both the design and implementation of these measures, political communication should be balanced and consider the extent to which resources available can be more effectively allocated to meet the needs of targeted households. Failure to do so can undermine the political acceptability of the measures and result in further discontent among underprivileged social groups. Together, these findings highlight the urgency of effective measures to contrast the rising phenomenon of poverty in Italy as characterised by widening geographical inequalities.

The takeaways from this dissertation can be summed up as follows. In a context of a deep regional divide and geographical inequalities, the paradigm of a place-based approach to regional development policies as developed by Barca (2009) is essential in order to overcome underdevelopment traps and tackle social exclusion. These goals are enshrined in the UN 2030 Agenda for Sustainable Development whose core principles are to leave no one behind and ensure that development targets are met for all individuals. For it to succeed, place-based strategies must facilitate partnerships between different levels of governance and recognise the importance of local knowledge to guide interventions. This approach hinge upon the availability of robust and timely evidence of socio-economic phenomena described at the granular level of detail. This dissertation is a testament to the potential that a local approach to social statistics can yield despite the limitations related to access to data. More pervasive and systematised use of administrative records at the micro-data level should represent the minimum condition for future academic and policy research on these issues. In addition, as highlighted in this dissertation, the challenges and possibilities ushered with the rise of the big data era call for the production of socio-economic data that are i) obtained from the integration of different thematic domains, ii) easily accessible and allow for the application of model inference methods and iii) granular in order to enable targeting, monitoring and policy evaluation. Based on these considerations, future lines of research could benefit from access to administrative microdata in different ways. In relation to the study of in-work poverty and sectoral employment, access to geo-localised microdata on earnings and employment would allow unpacking the findings observed around the role that sectors traditionally associated with low productivity and low wages play against a backdrop of de-industrialisation and economic contraction as present in the Southern Italian context. Further studies of the targeting of the RdC should assess the extent to which the measure has been effective in providing a safety net for the unprecedented loss of income as the results of the lockdown imposed to contain the spread of COVID-19, by combining data on RdC beneficiaries with tax returns registers. Finally, concerning Chapter 4 it is important to keep tracking the political evolution of the measure since a newly-elected right-wing government has announced to be restricting the pool of beneficiaries. Based on newly available data following the announced changes, further research should to study the response by affected voters as soon as the next European elections.

Bibliography

Alleva, Giorgio (2018). Relazione di fine mandato 2014-2018. Technical report, ISTAT.

Angrist, Joshua D and Jörn-Steffen Pischke (2014). *Mastering'metrics: The path from cause to effect*. Princeton university press.

Atkinson, Anthony B , Eric Marlier, Fabienne Montaigne, and Anne Reinstadler (2010). Income poverty and income inequality. *Income and living conditions in Europe 101*.

Baldini, Massimo , Paolo Bosi, and Stefano Tosi (2002). Targeting welfare in italy: old problems and perspectives on reform. *Fiscal Studies 23*(1), 51–75.

Baldini, Massimo , Giovanni Gallo, Lorenzo Lusignoli, Stefano Toso, et al. (2019). Le politiche per l'assistenza: il reddito di cittadinanza. Technical report, University of Modena and Reggio Emilia, Department of Economics "Marco Biagi".

Baldini, Massimo and Cristiano Gori (2019). Il reddito di cittadinanza. *il Mulino 68*(2), 269–277.

Barbieri, Paolo , Giorgio Cutuli, and Stefani Scherer (2018). In-work poverty in southern europe: The case of italy. In *Handbook on in-work poverty*. Edward Elgar Publishing.

Barca, Fabrizio (2009). *Agenda for a reformed cohesion policy*. European Communities Brussels.

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

Bartha, Attila , Zsolt Boda, and Dorottya Szikra (2020). When populist leaders govern: Conceptualising populism in policy making. *Politics and Governance 8*(3), 71–81.

Benavent, Roberto and Domingo Morales (2016). Multivariate fay–herriot models for small area estimation. *Computational Statistics & Data Analysis 94*, 372–390.

Benito, Juan Miguel and Roberto Ezcurra (2005). Spatial disparities in productivity and industry mix: the case of the european regions. *European Urban and Regional Studies 12*(2), 177–194.

Bennett, Fran (2013). Researching within-household distribution: Overview, developments, debates, and methodological challenges. *Journal of Marriage and Family 75*(3), 582–597.

Berlingieri, Giuseppe , Patrick Blanchenay, and Chiara Criscuolo (2017). The great divergence (s). *OECD Science, Technology and Industry Policy Papers* (39).

Bivens, Josh , Elise Gould, E Mishel, and Heidi Shierholz (2014). Raising america's pay. *Economic Policy Institute Briefing Paper 378*.

Bivens, Josh and Lawrence Mishel (2015). Understanding the historic divergence between productivity and a typical worker's pay: Why it matters and why it's real. *Economic Policy Institute. https://www. epi. org/files/2015/understanding-productivity-pay-divergence-final. pdf*.

Boeri, Tito (2019a). Audizione presidenza senato roma. Technical report, Istituto Nazionale di Previdenza Sociale.

Boeri, Tito (2019b). Dentro il reddito una norma iniqua - reddito, l'iniquità nascosta. *La Repubblica*.

Brandolini, Andrea (2021). Il dibattito sulla povertà in italia, tra statistica e politica (the poverty debate in italy: From politics to statistics). *Bank of Italy Occasional Paper* (648).

Burgard, Jan Pablo , Ralf Münnich, and Thomas Zimmermann (2015). Impact of sampling designs in small area estimation with applications to poverty measurement. *Analysis of Poverty Data by Small Area Estimation*, 83–108.

Callaway, Brantly , Andrew Goodman-Bacon, and Pedro HC Sant'Anna (2021). Difference-in-differences with a continuous treatment. *arXiv preprint arXiv:2107.02637*.

Casado-Díaz, José M , Mike Coombes, et al. (2011). The delineation of 21st century local labour market areas: a critical review and a research agenda. *Boletín de la Asociación de Geógrafos españoles*.

Casas-Cordero Valencia, Carolina , Jenny Encina, and Partha Lahiri (2016). Poverty mapping for the chilean comunas. *Analysis of Poverty Data by Small Area Estimation*, 379–404.

Cattaneo, Istituto (2018). Analisi della campagna elettorale, elezioni 2018. Technical report.

CEE (Commission of the European Communities) (1995). Final report on the implementation of the community programme concerning the economic and social integration of the economically and socially less privileged groups in society 'poverty 3' (1989-94). Technical report.

Checchi, Daniele , Giuseppe Pio Dachille, and Maria De Paola (2021). Reddito di cittadinanza, caratteristiche socio-economiche e capitale sociale (citizens' income, socio-economic characteristics, and social capital). *Politica economica 37*(1), 121–148.

Coady, David , Margaret Grosh, and John Hoddinott (2004a). Targeting outcomes redux. *The World Bank Research Observer 19*(1), 61–85.

Coady, David , Margaret E Grosh, and John Hoddinott (2004b). Targeting of transfers in developing countries: Review of lessons and experience.

Colantone, Italo , Gianmarco IP Ottaviano, and Piero Stanig (2021). The backlash of globalization.

Connelly, Roxanne , Christopher J Playford, Vernon Gayle, and Chris Dibben (2016). The role of administrative data in the big data revolution in social science research. *Social science research 59*, 1–12.

Cormier, David and Charles Craypo (2000). The working poor and the working of american labour markets. *Cambridge Journal of Economics 24*(6), 691–708.

Coveri, Andrea and Mario Pianta (2022). Drivers of inequality: wages vs. profits in european industries. *Structural Change and Economic Dynamics 60*, 230–242.

Curci, Nicola , Giuseppe Grasso, Pasquale Recchia, Marco Savegnago, et al. (2020). *Anti-poverty measures in Italy: a microsimulation analysis*. Banca d'Italia Eurosistema.

Currie, Janet (2004). The take up of social benefits. Technical report, National Bureau of Economic Research.

Datta, Gauri Sankar , Malay Ghosh, R Steorts, and J Maples (2011). Bayesian benchmarking with applications to small area estimation. *Test 20*(3), 574–588.

Daveri, Francesco , Cecilia Jona-Lasinio, and Francesco Zollino (2005). Italy's decline: Getting the facts right [with discussion]. *Giornale degli economisti e annali di economia*, 365–421.

David, H and David Dorn (2013). The growth of low-skill service jobs and the polarization of the us labor market. *American economic review 103*(5), 1553–97.

DDL n.1018 (2019, 02). *DL 4/2019 Reddito di cittadinanza e pensioni*.

De La O, Ana 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 Philippis, Marta , Andrea Locatelli, Giulio Papini, and Roberto Torrini (2022). La crescita dell'economia italiana e il divario nord-sud: trend storici e prospettive alla luce dei recenti scenari demografici [italian economic growth and the north-south gap: Historical trends and future projections in light of the recent demographic scenarios]. *Bank of Italy Occasional Paper* (683).

Dijkstra, Lewis , Hugo Poelman, and Andrés Rodríguez-Pose (2020). The geography of eu discontent. *Regional Studies 54*(6), 737–753.

DIPE (2022). Principali indicatori economici dell'economia italiana dal 2000.

D'ippoliti, Carlo and Alessandro Roncaglia (2011). L'italia: una crisi nella crisi (italy: a crisis within the crisis). *Moneta e Credito 64*(255), 189–227.

DLn.4/2019 (2019, 01). *Disposizioni urgenti in materia di reddito di cittadinanza e di pensioni*.

Dube, Arindrajit and Ethan Kaplan (2010). Does outsourcing reduce wages in the low-wage service occupations? evidence from janitors and guards. *ILR Review 63*(2), 287–306.

Duflo, Esther (2003). Grandmothers and granddaughters: old-age pensions and intrahousehold allocation in south africa. *The World Bank Economic Review 17*(1), 1–25.

Dwyer, Rachel E (2013). The care economy? gender, economic restructuring, and job polarization in the us labor market. *American Sociological Review 78*(3), 390–416.

Edmiston, Daniel (2022). Plumbing the depths: the changing (socio-demographic) profile of uk poverty. *Journal of Social Policy 51*(2), 385–411.

Einav, Liran and Jonathan Levin (2014a). The data revolution and economic analysis. *Innovation Policy and the Economy 14*(1), 1–24.

Einav, Liran and Jonathan Levin (2014b). Economics in the age of big data. *Science 346*(6210), 1243089.

El-Taliawi, Ola G , Nihit Goyal, and Michael Howlett (2021). Holding out the promise of lasswell's dream: Big data analytics in public policy research and teaching. *Review of Policy Research 38*(6), 640–660.

EUROSTAT (2021). Unemployment by sex and age (1992-2020) - quarterly data.

EUROSTAT (2022a). Living conditions in europe - work intensity, 2020. Technical report, EUROSTAT.

EUROSTAT (2022b). Real gdp growth rate - volume.

Faggio, Guilia and Stephen Nickell (2005). The responsiveness of wages to labour market conditions in the uk. *Labour Economics 12*(5), 685–696.

Farmer, Carson JQ and A Stewart Fotheringham (2011). Network-based functional regions. *Environment and Planning A 43*(11), 2723–2741.

Fay III, Robert E and Roger A Herriot (1979). Estimates of income for small places: an application of james-stein procedures to census data. *Journal of the American Statistical Association 74*(366a), 269–277.

Fierro, Paola (2019, Aug). Lettera al dirretore pasquale tridico. oggetto: sospensione dell'esame delle domande di reddito di cittadinanza presentate da cittadini stranieri.

Filandri, Marianna and Emanuela Struffolino (2019). Individual and household in-work poverty in europe: understanding the role of labor market characteristics. *European Societies 21*(1), 130–157.

Finan, Federico and Laura Schetcher (2012). Vote-buying and reciprocity. *Econometrica 80*(2), 863–881.

Foster-McGregor, Neil , Doris Hanzl-Weiss, Sandra M Leitner, Sebastian Leitner, Nirina Rabemiafara, Fadila Sanoussi, Robert Stehrer, and Terry Ward (2012). Sectoral employment effects of economic downturns. Technical report, wiiw Research Report.

Garnero, Andrea , Silvia Ciucciovino, Romolo de Camillis, Mariella Magnani, Paolo Naticchioni, Michele Raitano, Stefani Schrerer, and Emanuela Struffolino (2021). Relazione del gruppo di lavoro sugli interventi e le misure di contrasto alla poverta' lavorativa in italia. Technical report, Ministero del Lavoro.

Golden, Miriam and Brian Min (2013). Distributive politics around the world. *Annual Review of Political Science 16*, 73–99.

González-Manteiga, Wenceslao , María José Lombardía, Isabel Molina, Domingo Morales, and Laureano Santamaría (2008). Analytic and bootstrap approximations of prediction errors under a multivariate fay–herriot model. *Computational Statistics & Data Analysis 52*(12), 5242–5252.

Gori, Cristiano (2020). *Combattere la povertà: L'Italia dalla Social card al Covid-19*. Gius. Laterza & Figli Spa.

Gouldner, Alvin (1960). The norm of reciprocity: a preliminar statement. *American Sociological Review 25*, 161–178.

Gräbner, Claudius and Jakob Hafele (2020). The emergence of core-periphery structures in the european union: A complexity perspective. Technical report, ZOE Discussion Papers.

Gregg, Paul and Stephen Machin (2012). What a drag: the chilling impact of unemployment on real wages. *September.(Resolution Foundation.)*.

Gregg, Paul , Stephen Machin, and Mariña Fernández-Salgado (2014). Real wages and unemployment in the big squeeze. *The Economic Journal 124*(576), 408–432.

Gromadzki, Jan , Katarzyna Sałach, and Michal Brzezinski (2022). When populists deliver on their promises: the electoral effects of a large cash transfer program in poland. *Available at SSRN 4013558*.

Hadam, Sandra , Nora Würz, and Ann-Kristin Kreutzmann (2020). Estimating regional unemployment with mobile network data for functional urban areas in germany.

Hagenaars, Aldi JM , Klaas De Vos, M Asghar Zaidi, et al. (1994). Poverty statistics in the late 1980s: Research based on micro-data.

Hanna, Rema and Benjamin A Olken (2018). Universal basic incomes versus targeted transfers: Anti-poverty programs in developing countries. *Journal of Economic Perspectives 32*(4), 201–26.

Hicken, Allen (2011). Clientelism. *Annual Review of Political Science 14*(2), 289–310.

Hirschman, Albert O (1958). The strategy of economic development. Technical report.

Iammarino, Simona , Andrés Rodriguez-Pose, and Michael Storper (2019). Regional inequality in europe: evidence, theory and policy implications. *Journal of economic geography 19*(2), 273–298.

Imai, Kosuke , Gary King, and Carlos Velasco Rivera (2020). Do nonpartisan programmatic policies have partisan electoral effects? evidence from two large-scale experiments. *The Journal of Politics 82*(2), 714–730.

INPS (2019). Reddito/pensione di cittdaninanza e reddito di inclusione. osservatorio statistico. Technical report, Istituto Nazionale Previdenza Sociale.

Isserman, Andrew , Carol Taylor, Shelby Gerking, and Uwe Schubert (1987). Regional labor market analysis. In *Handbook of regional and urban economics*, Volume 1, pp. 543–580. Elsevier.

ISTAT (2014). La nuova geografia dei sistemi locali. Technical report.

ISTAT (2019). Le statistiche dell'istat sulla povertà. anno 2018. Technical report, Istituto Nazionale di Statistica.

ISTAT (2022a). Poverta' nuove serie.

ISTAT (2022b). Ricerca metodologica.

Jenkins, Stephen P (2020). Perspectives on poverty in europe. following in tony atkinson's footsteps. *Italian Economic Journal 6*(1), 129–155.

Jensen, Johan Ludwig William Valdemar et al. (1906). Sur les fonctions convexes et les inégalités entre les valeurs moyennes. *Acta mathematica 30*, 175–193.

Jiang, Jiming , P Lahiri, Shu-Mei Wan, and Chien-Hua Wu (2001). Jackknifing in the fayherriot model with an example. *Proc. Sem. Funding Opportunity in Survey Research*, 75–97.

Kaduk, Anne , Katie Genadek, Erin L Kelly, and Phyllis Moen (2019). Involuntary vs. voluntary flexible work: Insights for scholars and stakeholders. *Community, Work & Family 22*(4), 412–442.

Karlsson, Charlie and Michael Olsson (2006). The identification of functional regions: theory, methods, and applications. *The annals of regional science 40*, 1–18.

Kreutzmann, Ann-Kristin , Sören Pannier, Natalia Rojas-Perilla, Timo Schmid, Matthias Templ, and Nikos Tzavidis (2019). The R package emdi for estimating and mapping regionally disaggregated indicators. *Journal of Statistical Software 91*(7), 1–33.

Lawrence, Robert Z (2016). Does productivity still determine worker compensation? domestic and international evidence. *THE US LABOR MARKET-QUESTIONS AND CHALLENGES FOR PUBLIC POLICY*, 42–62.

Lohmann, Henning (2008). The working poor in european welfare states: Empirical evidence from a multilevel perspective. *The Working Poor in Europe: Employment, Poverty and Globalization*, 47–74.

Lucidi, Federico and Alfred Kleinknecht (2010). Little innovation, many jobs: An econometric analysis of the italian labour productivity crisis. *Cambridge Journal of Economics 34*(3), 525–546.

Lundberg, Shelly and Robert A Pollak (1996). Bargaining and distribution in marriage. *Journal of economic perspectives 10*(4), 139–158.

Machin, Stephen (2016). Rising wage inequality, real wage stagnation and unions. In *Inequality: Causes and consequences*. Emerald Group Publishing Limited.

MacKinnon, Danny , Louise Kempton, Peter O'Brien, Emma Ormerod, Andy Pike, and John Tomaney (2022). Reframing urban and regional 'development' for 'left behind' places. *Cambridge Journal of Regions, Economy and Society 15*(1), 39–56.

Manacorda, Marco , Edward Miguel, and Andrea Vigorito (2011). Government transfers and political support. *American Economic Journal: Applied Economics 3*(3), 1–28.

Manasse, Paolo (2013). The roots of the italian stagnation. *CEPR Policy Insight 15*(66), 2013.

McCann, Philip (2020). Perceptions of regional inequality and the geography of discontent: Insights from the uk. *Regional Studies 54*(2), 256–267.

Meyer, Bruce D , Wallace KC Mok, and James X Sullivan (2015). Household surveys in crisis. *Journal of Economic Perspectives 29*(4), 199–226.

Meyer, Bruce D and James X Sullivan (2003). Measuring the well-being of the poor using income and consumption. Technical report, National Bureau of Economic Research.

Mishel, Lawrence (2018). Yes, manufacturing still provides a pay advantage, but staffing firm outsourcing is eroding it.

Monducci, Roberto (2019). Esame del disegno di legge n. 1080. "conversione in legge del decreto-legge 28 gennaio 2019 n.4, recante disposizioni urgenti in materia di reddito di cittadinanza e di pensioni". Technical report, ISTAT.

Natale, Paolo (2014). The birth, early history and explosive growth of the five star movement. *Contemporary Italian Politics 6*(1), 16–36.

OECD.Stat (2022). Average annual wages.

O'Connor, Alice (2016). Poverty knowledge and the history of poverty research. *The Oxford handbook of the social science of poverty*, 169–192.

Pedrazzani, Andrea and Luca Pinto (2017). Dove pesca la rete del movimento: le basi sociali del suo voto.

Permatasari, Novia and Azka Ubaidillah (2021). *msae: Multivariate Fay Herriot Models for Small Area Estimation*. R package version 0.1.4.

Pinheiro, Flavio L , Pierre-Alexandre Balland, Ron Boschma, and Dominik Hartmann (2022). The dark side of the geography of innovation: relatedness, complexity and regional inequality in europe. *Regional Studies*, 1–16.

Pratesi, M (2017). I big data: il punto di vista di uno statistico. *Il menabò-Associazione Etica ed Economia, Menabò* (62).

Pratesi, Monica and Nicola Salvati (2016). Introduction on measuring poverty at local level using small area estimation methods. *Analysis of poverty data by small area estimation*, 1–18.

Rao, John NK and Isabel Molina (2015). *Small area estimation*. John Wiley & Sons.

Ravallion, Martin (2009). How relevant is targeting to the success of an antipoverty program? *The World Bank Research Observer 24*(2), 205–231.

Rodríguez-Pose, Andrés (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.

Saraceno, Chiara (2015). *Il lavoro non basta: la povertà in Europa negli anni della crisi*. Feltrinelli Editore.

Saraceno, Chiara (2020). *Quando avere un lavoro non basta a proteggere dalla povertà*. Firenze: Firenze University Press.

Saraceno, Chiara and David Benassi (2020). *Poverty in Italy: Features and Drivers in a European Perspective*. Policy Press.

Saraceno, Chiara , David Benassi, and Enrica Morlicchio (2022). *La povertà in Italia: soggetti, meccanismi, politiche*. Il Mulino.

Satterthwaite, David and Cecilia Tacoli (2002). Seeking an understanding of poverty that recognizes rural-urban differences and rural-urban linkages. *Urban Livelihoods: A people-centred approach to reducing poverty 52*, 70.

Sawicky, Max B (2002). The mirage of welfare reform. *WorkingUSA 6*(3), 55–69.

Schmid, Timo , Fabian Bruckschen, Nicola Salvati, and Till Zbiranski (2017). Constructing sociodemographic indicators for national statistical institutes by using mobile phone data: estimating literacy rates in senegal. *Journal of the Royal Statistical Society: Series A (Statistics in Society) 180*(4), 1163–1190.

Schwellnus, Cyrille , Andreas Kappeler, and Pierre-Alain Pionnier (2017). Decoupling of wages from productivity: Macro-level facts.

Shucksmith, Mark (2003). Social exclusion in rural areas: a review of recent research. *Arkleton Centre for Rural Development Research, University of Aberdeen, Aberdeen*.

Sissons, Paul , Anne E Green, and Neil Lee (2018). Linking the sectoral employment structure and household poverty in the united kingdom. *Work, Employment and Society 32*(6), 1078–1098.

Snel, Erik , Jan de Boom, and Godfried Engbersen (2008). The silent transformation of the dutch welfare state and the rise of in-work poverty. *The working poor in Europe: employment, poverty and globalization*, 124–154.

Stansbury, Anna M and Lawrence H Summers (2018). Productivity and pay: Is the link broken? Technical report, National Bureau of Economic Research.

Strain, Michael R (2019). The link between wages and productivity is strong. *Expanding Economic Opportunity for More Americans, The Aspen Institute*.

Szirmai, Adam (2012). Industrialisation as an engine of growth in developing countries, 1950–2005. *Structural change and economic dynamics 23*(4), 406–420.

Tanton, Robert , ANN Harding, and Justine McNamara (2010). Urban and rural estimates of poverty: Recent advances in spatial microsimulation in australia. *Geographical Research 48*(1), 52–64.

The Economist (2005). The real sick man of europe.

Tomlinson, Mark and Robert Walker (2010). Recurrent poverty: the impact of family and labour market changes. *York: Joseph Rowntree Foundation*.

Torrini, Roberto , Giuseppe Albanese, Antonio Accetturo, Domenico Depalo, Valerio Paolo Vacca, Giovanna Messina, Filippo Scoccianti, and Silvia Giacomelli (2022). Il divario nord sud: sviluppo economico e intervento pubblico. In *Bank of Italy Workshops and Conferences series*.

Tronconi, Filippo (2015). *Beppe Grillo's Five Star Movement: Organisation, Communication and Ideology*. Ashgate Publishing, Ltd.

Tzavidis, Nikolaos , Li-Chun Zhang, Angela Luna Hernandez, Timo Schmid, and Natalia Rojas-Perilla (2018). From start to finish: a framework for the production of small area official statistics. *Journal of the Royal Statistical Society: Series A (Statistics in Society) 181*(4), 927–979.

Union, European , UN-Habitat, OECD, and World Bank (2021). Applying the degree of urbanisation: A methodological manual to define cities, towns and rural areas for international comparisons (2021 edition).

Vannutelli, Silvia (2023). The political economy of stimulus transfers.

Vanselow, Achim , Chris Warhurst, Annette Bernhardt, and Laura Dresser (2010). Working at the wage floor: hotel room attendants and labor market institutions in europe and the united states. *Low-wage work in the wealthy world*, 269–318.

Visaria, Pravin (1980). Poverty and living standards in asia. *Population and Development Review*, 189–223.

Wang, Man , Rachel Garshick Kleit, Jane Cover, and Christopher S Fowler (2012). Spatial variations in us poverty: Beyond metropolitan and non-metropolitan. *Urban Studies 49*(3), 563–585.

Wannell, Ted and Jeannine Usalcas (2012). Labour force survey: 2011 year-end review. *Perspectives on Labour and Income 24*(2), 1.

Weziak-Bialowolska, Dorota (2016). Spatial variation in eu poverty with respect to health, education and living standards. *Social indicators research 125*(2), 451–479.

White, Patricia and R Saylor Breckenridge (2014). Trade-offs, limitations, and promises of big data in social science research. *Review of Policy Research 31*(4), 331–338.

Wolter, Kirk (2007). *Introduction to variance estimation*. Springer Science & Business Media.

Zucco Jr, Cesar (2013). When payouts pay off: Conditional cash transfers and voting behavior in brazil 2002–10. *American journal of political science 57*(4), 810–822.

Statement of contribution and dissemination

All chapters, with the exception of Chapter 1, were made in collaboration with the other co-authors. The specific contributions of all co-authors to the contents are detailed below.

- **Chapter 2:** Contributions to the chapter come from the authors Giovanni Tonutti, Gaia Bertarelli, Andrea Garner and Monica Pratesi. G. Tonutti and A. Garnero jointly developed the research idea, defined the indicators and worked on the research design. M. Pratesi provided access to both the EUSILC data with geo-referenced information and the ASIA register. G. Tonutti cleaned the data, produced the SAE estimates and carried out the regression analysis. G. Bertarelli reviewed the production of the estimates and contributed to drafting the paper.

- **Chapter 3:** Contributions to the chapter come from the authors Giovanni Tonutti, Gaia Bertarelli, Caterina Giusti and Monica Pratesi. The initial research idea stems from G. Tonutti. M. Pratesi and C. Giusti provided mentorship and access to EU-SILC and HBS data . G. Bertarelli was responsible for the selection of the estimation methods and contributed to drafting of the paper.

  The paper has been presented at the 63rd ISI World Statistics Congress 2021 (virtual), at the InGrid2 Workshop on 'Measuring and monitoring regional multidimensional poverty and cost of living: new data, new methods' in March 2021 and at the IES 2022 Conference: 'Statistical and Economics Methodology for Quality Assessment' at the Universita' L. Vanvitelli, Capua (CE) in January 2022. The work has been published on the journal *Socio-Economic Planning Sciences, Volume 82. August 2022*.

- **Chapter 4:** Contributions to the chapter come from the authors Giovanni Tonutti and Mauro Sylos-Labini. The initial research idea stems from M. Sylos-Labini. G. Tonutti was responsible for cleaning and harmonizing the election data and carrying out the regression analysis, in collaboration with M. Sylos-Labini. M. Sylos-Labini reviewed the version of the paper included in this thesis.