UNIVERSITÀ COMMERCIALE “LUIGI BOCCONI”

PhD SCHOOL

PhD program in Public Policy and Administration

Cycle: XXXIII

Disciplinary field (code): SECS-S/04

Essays in Population Health

Advisor: Simone GHISLANDI

PhD Thesis by

Benedetta SCOTTI

ID number: 1648856

Year 2022

Acknowledgments

I want to express my deepest gratitude to all those who made this thesis possible. I am extremely grateful to Prof. Simone Ghislandi who supervised this work with patience and generosity, giving me the possibility to work together on the Visitinps project. Likewise, I want to thank Prof. Franceso C. Billari and Prof. Alessia Melegaro who accepted to participate to the thesis committee, providing useful comments and suggestions. I am also grateful to Prof. Raya Muttarak and Dr. Markus Sauerberg, whom I had the chance to work with in the past years and who contributed to the research material of the fourth chapter of this thesis.

PhD life would have been much less stimulating without the fellowship of Rita, Selin, Chiara, Paula, Francesco, Samir, Silvia, Davide. You made our long days in the room 3H brighter.

I owe an immense debt of gratitude to my family. First of all, I want to thank my parents, Andrea and Alessandria, for the unconditional support they provided me with throughout all these years, from a material, moral and spiritual standpoint. I am grateful to my brothers, Giovanni and Luca, for their lovely goodwill. A note on Luca's readings on Central Asia: your obsession for Turkmenistan made the lockdown less gloomy. I want also to express my gratitude to my aunt Maria Chiara and my uncle Massimo, for the great generosity they have always shown since my very first arrival in Milan.

I want to thank all the friends who managed to take me away from the thesis, offering me moments of joy and hope. A huge thank to Martina, who forgave my daily faults and my culinary inability, making all trips to Milan merrier; to all my friends of the Little Oratory of Rome, in particular to Padre Maurizio, Chiara, Clotilde, Maria Amalia, Benedetta and Gianfranco, for always welcoming me back, and to Francesca and Olimpia, with whom I shared great moments of PhD life in the library of Palazzo Farnese; and to Marco and Laura Invernizzi, who opened me the doors of their home a few steps from Bocconi.

This thesis would have not been possible without the love, the encouragement and the kind patience of Mirko, who asked me to marry him, despite myself. Thanks for all the times you brought me back to reality.

And last, but certainly not least, thanks to the good Lord, who is the beginning and the end of everything.

A.M.D.G.

Rome, 22 November 2021

Ringraziamenti

Non avrei potuto portare a termine gli studi di dottorato senza il supporto di professori, familiari e amici che, a vario titolo, mi hanno accompagnato in questi quattro anni di gioie e di fatiche. Il mio primo caloroso e sincero ringraziamento va al Prof. Simone Ghislandi che ha supervisionato con pazienza e generosità questa tesi, offrendomi la possibilità di lavorare assieme al progetto Visitinps e incoraggiandomi nei momenti di maggior scoramento. Desidero anche esprimere la mia riconoscenza nei confronti del Prof. Francesco C. Billari e della Prof.ssa Alessia Melegaro, che hanno accettato di far parte della commissione di tesi e che hanno saputo darmi consigli e suggerimenti preziosi. Un sentito ringraziamento anche alla Prof.ssa Raya Muttarak e al Dott. Markus Sauerberg con cui ho avuto il piacere di collaborare in questi anni e che hanno contribuito al materiale di ricerca che è confluito nel quarto capitolo di questa tesi.

La vita in università sarebbe stata certamente meno stimolate senza la compagnia dei miei colleghi di avventura dottorale: Rita, Selin, Chiara, Paula, Francesco, Samir, Silvia, Davide. A loro la mia gratitudine per aver reso più luminose le lunghe giornate di studio nella famigerata stanza 3H.

Ho un debito di gratitudine incolmabile nei confronti della mia famiglia, che mi ha sopportato con benevolenza e pazienza per tutti questi anni. Desidero ringraziare in primis i miei genitori, Andrea e Alessandra, che non mi hanno mai fatto mancare il loro supporto incondizionato, da un punto di vista materiale, morale, e spirituale, e i miei fratelli, Giovanni e Luca, che hanno saputo sempre mostrarmi il loro affetto, soprattutto nei momenti di scoramento. Una particolare nota di merito a Luca: senza le sue letture sull'Asia Centrale, i mesi di stesura della tesi durante il confinamento a casa a causa della pandemia sarebbero stati terribilmente grigi. Un grande ringraziamento anche ai miei zii, Maria Chiara e Massimo, che con la loro generosità hanno saputo addolcire fin dal principio la permanenza milanese lontano da Roma.

Desidero esprimere tutta la mia gratitudine nei confronti degli amici che con grande caparbietà hanno saputo allontanarmi da libri, articoli e computer, regalandomi momenti di gioia e speranza. Un grazie immenso a Martina, che ha saputo perdonare le mie mancanze quotidiane e colmare le mie lacune culinarie, rendendo felice ogni salita a Milano; a tutti gli amici dell'Oratorio Piccolo di Roma, in particolare a Padre Maurizio, Chiara, Clotilde, Maria Amalia, Benedetta e Gianfranco, che mi hanno sempre accolto a braccia aperte ad ogni ritorno a Roma, e a Olimpia e Francesca con cui ho condiviso le amenità della vita da dottoranda e della biblioteca di Palazzo Farnese; a Marco e Laura Invernizzi, che mi hanno fatto scoprire un porto amico a pochi passi dalla Bocconi.

Il lavoro di tesi sarebbe stato infinitamente più faticoso senza l'affetto, l'incoraggiamento e la paziente gentilezza di Mirko che mi ha chiesto di diventare sua moglie nonostante tutto. Grazie per tutte le volte in cui sei stato il mio ritorno al reale.

E infine, grazie al buon Dio, che di ogni cosa è principio e compimento.

A.M.D.G.

Roma, 21 Novembre 2021

“Quis autem vestrum cogitans

potest adicere ad aetatem suam cubitum unum?”

Evangelium secundum Matthaeum 6, 27

Thesis abstract

Since the early 21st century, concerns about widening inequalities in social and economic outcomes have come to dominate academic and political debates. As noted by van Raalte et al. (2018), the first and most fundamental form of inequality is inequality in lifespan as all other types of inequality are conditional on being alive. Studying and tackling the root causes of mortality inequalities is thus key for unveiling, analysing and addressing socio-economic disparities along other dimensions. Moreover, besides being troubling *per se*, mortality inequalities have also profound implications when it comes to the set-up of key policy programs ranging from public healthcare to social security systems. Building on these insights, in this thesis I investigate how mortality inequalities are shaped by life-course socio-economic circumstances and by exogenous demographic shocks. The thesis is articulated in four chapters sharing the same empirical context, i.e. Italy, and characterized by extensive reliance on administrative data. The first three chapters, based on data extracted from the digital archives of the Italian Social Security Institute (INPS), provide novel evidence about mortality inequalities arising from employment trajectories, occupational background and socio-economic status among the adult and elderly population in Italy, and about their policy implications for the public pension system. The last chapter, which builds upon a compendium of data from the Italian National Institute of Statistics (ISTAT) including newly digitized archival sources, investigates and compares spatial inequalities in the mortality impact of major influenza outbreaks since the early 1950s up to the COVID-19 pandemic at the subnational level in Italy.

In Chapter 1, I analyse how employment instability in later working life shapes post-retirement survival chances in a large sample of retirees previously employed in the private sector in Italy. While the consequences of discontinuous employment trajectories on health and well-being of young and middle-aged people have been extensively studied, little attention has been devoted to the relationship between employment instability in later working life and subsequent health outcomes. Using data retrieved from the Italian Social Security Institute archives, this chapter provides empirical evidence addressing this gap by analysing the relationship between late-career trajectories and post-retirement mortality among Italian retirees. Methodologically, I employ sequence analysis to identify ideal-type late-career trajectories, that I use as explanatory variables in discrete-time survival analysis. Results suggest that going through periods of unemployment prior to retirement, especially if prolonged and not covered by unemployment allowances, is negatively linked to post-retirement survival. Given the current pressures to extend the length of working life, these findings highlight the relevance of policies aimed at improving the employment prospects of displaced senior workers, and at ensuring adequate support in case of prolonged unemployment.

In Chapter 2, I examine all-cause mortality differentials among Italian retirees by former occupational class, defined on the basis of a highly detailed taxonomy encompassing over thirty categories. The analysis presented in this chapter tackles a major limitation of existing studies on occupation-based mortality inequalities, i.e. their reliance on broadly defined occupational classes, which limits their relevance and usefulness for policy purposes. I find mortality patterns among Italian retirees by former occupation to be markedly gendered. In the case of men, I observe increasing mortality risk moving from highly qualified, non-manual occupations (such as engineers and architects) to manual, low-skilled and labour-intensive occupations (such as labourers in mining, construction and manufacturing). I estimate that at 65 male retirees with a background in specific low-risk occupational categories enjoy an advantage in residual life expectancy of about 4-5 years compared to those with a background in specific high-risk categories. I document limited occupation-based mortality gradient, instead, among female retirees. In fact, women with a background in specific upper non-manual and high-skilled occupations face worse survival chances into old-age compared to females previously employed in lower manual and elementary occupations. The results of this study are highly relevant for policymakers in Italy, and in countries confronting the need of extended working life policies to meet rising pressures on social security systems posed by increased longevity. In particular, they highlight the importance of policy measures aimed at ensuring flexible retirement options to individuals belonging to specific high-risk occupational groups.

In Chapter 3, I explore the evolution of lifespan inequalities among the over 50 in Italy, evaluating the distributional implications of such inequalities for the Italian pension system. Previous research has documented the existence of tangible disparities in longevity by socio-economic status in Italy, especially among men, discussing the challenges they pose to the equity and the sustainability of the pension system (Ardito et al. 2021). This literature presents, though, three main limitations. First, it is overwhelmingly cross-sectional in nature. This gap is particularly relevant in the light of the growing number of studies documenting widening inequalities in longevity by socio-economic status in several OECD countries (Sasson 2016). Second, it focuses exclusively on life expectancy (average age-at-death) differentials. Still, as suggested by a growing strand of literature, monitoring both *average* age-at-death and *variation* in age-at-death is crucial for gaining full insight about heterogeneity in population health (van Raalte et al. 2018). Third, discussions about the distributional implications of differential mortality for the Italian pension system are not supported by evidence based on real employment and contributory histories. As such, little is known about the real degree of redistribution implied by heterogeneity in longevity in the Italian pension system. The analyses presented in this chapter address all of these points. First, I document how life expectancy and lifespan variation by socio-economic status, measured in terms of lifetime income, have evolved over birth cohorts (1930-1957) and calendar years (1995-2017) among individuals with an employment background in the private sector in Italy. Results shed light on some worrisome trends in the evolution of lifespan inequalities among the over 50 in Italy. Although mortality delay (increasing average age-at-death) and mortality compression (declining lifespan variability) are observed across all socio-economic strata, such improvements were reaped mostly at the top of the lifetime income distribution, notably in the case of men. In the case of women, lifespan inequalities are limited, with no clearly discernible trends over time. I then show that the distributional implications of such inequalities in lifespan for the pension system are tangible. In particular, I document that the erosion in the profitability of pension contributions implied by heterogeneous longevity is stronger for male retirees at the bottom of the lifetime income distribution, and that such dynamics have become more pronounced over time.

Finally, in Chapter 4, I compare the provisional death toll of COVID-19 in Italy against the mortality burden of major flu pandemics in the post-World War II period, i.e. the 1957 and 1969 flu pandemics, accounting for spatial inequalities in mortality at the sub-national level. For this purpose, I exploit a newly constructed digital database with age-sex-specific yearly mortality and population patterns in Italian provinces covering nearly seven decades (1953-2020). Methodologically, I estimate all-cause excess mortality in Italian provinces in pandemic years, quantifying the impact on life expectancy and lifespan variation. I find that in peak years 1957 and 1969, life expectancy at birth for women and men in hardest hit Italian provinces was respectively 1.5-2 and 2-3.1 years lower than expected. I also document that in hardest hit areas lifespan variation was higher than expected in both 1957 and 1969 as a consequence of increased infant mortality. Compared to these major flu events, I document that the human costs of the COVID-19 pandemic in 2020 have been substantially larger. In 2020, in most affected provinces life expectancy for men and women dropped by 4.6 and 3 years respectively. As a result of the sharp rise in mortality at older ages, in 2020 lifespan variation decreased too. To the best of my knowledge, this study represents the first direct comparison of the mortality impact of the COVID-19 pandemic at the subnational level with the main past flu events experienced since the end of World War II.

Contents

Chapter 1

Late-career employment trajectories and post-retirement mortality: evidence from Italy

1.1 Introduction

The relationship between employment and retirement has changed dramatically over the last decades in advanced economies. The early life-course view which postulated predictable career paths, with lifetime employment ending with direct entry into full retirement at a specific age, has been challenged by evidence that retirement patterns have become increasingly diversified (Fasang 2010; 2012). Behind the general trend towards early retirement observed throughout the 1970-1990s, and its progressive reversal through pension reforms aimed at postponing retirement, important heterogeneities have emerged. For many, leaving their major employment before being meeting retirement eligibility requirements is an involuntary choice. For others, it is a strategy to cope with policy-induced changes in the length of working life. As such, the period between the end of stable employment and entry into full retirement at pensionable age can be quite discontinuous, marked by frequent job searches, displacements and unemployment spells,

[1]The paper constituting this chapter is currently at the second round of review at *Demography.* I am grateful to F.C. Billari, C. Devillanova, S. Ghislandi, E. Struffolino, to partecipants to the DisCont research group at Bocconi University and to three anonymous referees for useful comments and suggestions. Access to INPS data was kindly granted by the Italian Ministry of Labour. The opinions expressed in this study are those of the author. They do not purport to reflect the opinions or views of the Italian Ministry of Labour or its members.

or re-employment at lower occupational status and salary (Marshall et al. 2001; Riekhoff & Järnefelt 2018). Rather than being a definite event in time, marking the passage from paid work to pension income, retirement looks increasingly as a process consisting in a "series of acts involving movements both out and back in the labor force" (Hayward & Grady 1990; p. 352).

Although there is an extensive literature examining the relationship between retirement and health, we know little about the consequences of discontinuous employment trajectories in later working life on health. This is a major gap given that a sizeable, possibly growing, proportion of older workers in advanced economies is now facing the prospect of some form of employment instability, encompassing job loss, prolonged unemployment or repeated transitions into and out of unemployment, before entering into full retirement. This study provides empirical evidence addressing this gap. I do so by identifying ideal-type late-career trajectories for a large sample of individuals who spent most of their professional life working as private employees in Italy, who retired between 2001 and 2018, and for whom I can observe mortality up to 2018. For this purpose, I use register data from the Italian Social Security Institute (INPS) which allow to accurately reconstruct individual working histories. Focusing on the ten years prior to retirement, I identify the prevalent forms of employment instability, which materialize as deviations from full-time private employment, relating them to post-retirement mortality. Methodologically, I employ sequence analysis to identify ideal-type late-career trajectories that I use as explanatory variables in discrete-time survival analysis. Results suggest that going through periods of unemployment prior to retirement, especially if prolonged and not covered by unemployment allowances, impacts negatively on post-retirement survival, the effect being stronger for men as compared to women. I also find that the adverse consequences of employment instability in later working life on survival tend to be more acute among disadvantaged occupational sub-groups, at least in the case of women. These results are robust to sensitivity analyses which aim at alleviating concerns on reverse causality dynamics between employment trajectories and health.

This work draws on and contributes to three main strands of literature. First, it provides novel evidence on the 'scarring effects' of employment instability (Clark et al. 2001), focusing on a period of working life, the later one, which has been largely neglected so far. To the best of my knowledge, this is the first study investigating how late-career employment trajectories relate to post-retirement survival chances. Second, it integrates the growing literature on the relationship between retirement and subsequent health outcomes. It does so by stressing the importance of taking a dynamic approach to retirement, in the wake of studies which look at retirement as a sequentially linked process rather than as a single transition in time (Fasang 2010; 2012; McDonough et al. 2017; Riekhoff 2018; Riekhoff & Järnefelt 2018). It is worth stressing that this study does not explore the relationship between unemployment or retirement and mortality *per se*, but rather the relationship between employment instability in the path to retirement and subsequent mortality. Last, this study speaks also to the well-established literature on differential mortality (Marmot 2005; Wilkinson & Marmot 2003). To the extent that deviations from full-time employment in later adult life are more frequent across socio-economically vulnerable strata of the population, it can shed further light on the relationship between socio-economic status and lifespan inequalities. Also, it can provide new insights on drivers of differential mortality within socio-economic groups themselves.

The remainder of this study is organized as follows. In Section 1.2, I provide a concise overview of the relevant literature which I contribute to. In Section 1.3 and 1.4, I present the data and the methodology I adopt. I then proceed with the results, that I illustrate in Section 1.5. Final remarks, including comments on limitations and avenues for improvements and further research, conclude.

1.2 Employment instability, retirement and health

The relationship between employment and health has long been studied (Dooley et al. 1996; Jin et al. 1995; Roelfs et al. 2011). Still, it continues to raise significant scholarly interest. Indeed, the mechanisms connecting employment status and health outcomes are complex, possibly countervailing and certainly difficult to disentangle. On the one hand, precarious employment, unemployment, and inactivity are generally expected to have a negative impact on individual health, the potential channels being, *inter alia*, economic hardship, loss of status, or harmful behavior, such as suicide (Wilkinson & Marmot 2003). On the other hand, employment itself can have a negative impact on health via physical stress, injuries, professional diseases, or perceived job insecurity (Caroli & Godard 2014; Quinlan 2015). The consensus is that unemployment is associated with worse health outcomes, also after accounting for health selection issues (Blakely et al. 2003; Pirani & Silvani 2015; Vågero & Garcy 2016). Yet, evidence of causality remains mixed (Salm 2009; Schmitz 2011; Sullivan & von Wachter 2009). A major limitation of many existing studies connecting employment and health, which may also be the reason behind contradictory results, is their focus on short-term labour market events. In fact, the effects of given employment conditions may depend on the length of exposure to a given status, accumulate over time and emerge only gradually. For this reason, scholars have increasingly stressed the need for taking a more dynamic approach to the issue, looking at long-run employment trajectories over the life course (Benach & Muntaner 2007). Longitudinal data, and techniques to process them, such as sequence analysis, have been employed for this purpose. Research efforts have concentrated mostly on the long-term impact of employment trajectories on health and well-being of young and middle-aged people (Clark & Lepinteur 2019; Devillanova et al. 2019; Sarti & Zella 2016; Torssander & Almquist 2017; Waenerlund et al. 2014). These studies are unanimous in concluding that low labour market attachment and trajectories marked by spells of short- and long-term unemployment have a negative impact on self-reported health and survival.

Little attention has been devoted, though, to the relationship between employment instability in later working life or in the retirement process and subsequent health outcomes. Research on transition into retirement and health has focused overwhelmingly on the effect of the transition *per se* or on the timing. The overall evidence on the effect of retirement on health is, at best, mixed (van der Heide et al. 2013). As for timing, the consensus is that early retirement has adverse effects on health (Burdorf 2010; Hult et al. 2010; McDonough et al. 2017; Westerlund et al. 2010), even after accounting for health selection issues (Barban et al. 2017). Some studies have examined the consequences of employment instability in the retirement process for income inequality in old age (Fasang 2012; Riekhoff & Järnefelt 2018). With specific reference to the Italian context, Contini & Leombruni (2006) find that the late-career of a non-negligible share of the working population in the early 1990s was marked by irregular patterns of labour market activity, with negative consequences for wages and pensions. Still, there is limited research about the relationship between employment-to-retirement trajectories and inequality in seniors' health and longevity. A few studies employ SHARE data providing retrospective employment information to investigate how employment histories over the life course relate to self-reported health and quality of life at older ages (Ponomarenko 2016; Wahrendorf 2015; Zella & Harper 2018). These studies suggest that going through spells of inactivity and unemployment over the life time has negative consequences for well-being in old age, particularly in the case of men. To the best of our knowledge, Marshall et al. (2001) provides the only study, so far, looking specifically at the relationship between late-career employment trajectories and subsequent health outcomes. Focusing on a sample of early retirees from a major Canadian telecommunication company in the 1985-1995 decade, they find that instability in the retirement transition, measured in terms of unemployment spells, yields adverse effects on self-rated health after retirement. Analogous evidence on how late-career employment trajectories relate to post-retirement mortality is still missing. Given the progressive erosion of direct transition from life-time employment to retirement, there is room for arguing that further research on late-career trajectories and health is much due.

1.3 Data

1.3.1 The INPS LoSai sample

I use longitudinal register data extracted from the archives of the Italian Social Security Institute (INPS) to examine late-career trajectories of a sample of workers previously employed in the Italian private sector, who retired between 2001 and 2018. Specifically, I rely on the INPS LoSai sample made available by the Italian Ministry of Labour.[1] The INPS LoSai sample is made up of individuals born on days 1 and 9 of each month of any birth year, for all cohorts appearing in the INPS archives. To the extent births are uniformly distributed in any year, extracted individuals represent about 8% of the INPS population. I combine data from three sources. First, the *Estratti Conto* dataset, which reports the entire contributory history of private sector workers covered by INPS-managed social security schemes until 2018. Major schemes managed by INPS cover private sector employees and self-employed workers, including craftsmen, shop-keepers, and farmers. Public employees and high-profile freelance professionals (e.g. architects, lawyers, etc.) are covered, instead, by non-INPS schemes.[2] The *Estratti Conto* dataset provides a detailed record of all episodes in one's working life covered by INPS social security contributions: employment/self-employment job spells, parental/family leaves, sickness/injury episodes, unemployment spells covered by social benefits, work suspension spells covered by wage subsidy schemes. For each spell, there is information about its beginning/ending date, and about the monetary value of the relative contributions. For

[1] Access to the INPS LoSai sample is granted to researchers upon the presentation of a research proposal. See `https://www.cliclavoro.gov.it/Barometro-Del-Lavoro/Pagine/Microdati-per-la-ricerca.aspx`

[2] Public employees' pension scheme (INPDAP) was brought under INPS' supervision at the end of 2011. Still, public employment spells are not available in the INPS LoSai sample.

private employment spells, the *Estratti Conto* provides also information about occupational class (blue-collar, white-collar, middle-manager and manager) and type of contract (full-time vs part-time). The second source I draw on is the *Casellario Pensioni* archive, which keeps track of all recipients of pension benefits disbursed by INPS between 2001 and 2018. For each pension benefit, I have information about the date in which the pension flow started, the date in which the pension flow ended (if this occurs by the end of 2018) and the type of pension benefit.[3] Appendix A.1 provides a detailed description of categories of pension benefits disbursed by INPS. Third, I use information extracted from the *Anagrafica* dataset, which reports basic demographic characteristics of workers appearing in the INPS archives, including gender, year of birth, year of death (if relevant), and region of residence as of 2018 or as of the year of death.[4] A major drawback of the *Anagrafica* dataset, and of register data in general, is the paucity of information about individual characteristics. I notably lack information about education, marital/family status, and other family background characteristics, and I cannot match across spouses or family members.

1.3.2 Sample construction

The initial sample is composed by 501,220 individuals who retired between 2001 and 2018. I identify retirement time as the first year in which the individual starts receiving old-age or seniority pension benefits from INPS.[5] I restrict my analysis to individuals who report at least one contributory spell (i.e. who appear in the *Estratti Conto* dataset) in the ten years prior to retirement and who spent most of their career working as private employees,

[3]The INPS LoSai sample allows to distinguish five types of pension benefits: (i) old-age/seniority pensions, which are granted upon attainment of age/contributory requirements; (ii) disability pensions; (iii) survivors' pensions; (iv) indemnity pensions, which are granted to individuals affected by occupational injuries/diseases; (v) social pensions, which encompass benefits to severely disabled people and to individuals with old-age/seniority pension benefits below the minimum. The monetary amount of pension benefits is not available in the INPS LoSai sample.

[4]Data in *Anagrafica* do not allow to identify past changes of residence, if any

[5]Individuals are entitled to old-age pension benefits upon attainment of specific age. Entitlement to seniority pension benefits is related, instead, to years of social security contributions. In the INPS LoSai sample, these two types of pension benefits are not distinguishable.

as deducible from their contributory history (N=238,098). In other words, I focus on individuals who contributed for most of their career, measured in terms of years (i.e. > 50%), to the INPS scheme covering private sector employees (*Fondo Pensione Lavoratori Dipendenti*).[6] I also exclude individuals whose first appearance in INPS archives occurs at a suspiciously young age, i.e. below 12, and individuals whose first appearance in INPS archives occurs at age > 30, the purpose being to focus on individuals whose career is mostly tracked by the *Estratti Conto* dataset (N=8,490). To alleviate reverse causality concerns about late-career trajectories and subsequent mortality patterns, I further exclude individuals who spend most of any of the ten years prior to retirement in injury/sickness leave (N=1,857). I also exclude individuals who retired before the age of 50 or after the age of 70 (N=3,253), the aim being to have a sample of people who entered into retirement at a "normal" age, presumably in good health. The final sample consists of 224,498 uniquely identified individuals, 146,978 men and 77,520 women[7], that can be followed from retirement up to 2018 or to the year of death if this occurs before 2018, for a total of 2,206,937 person-year observations and 20,379 recorded deaths (16,458 in the case of men, 3,921 in the case of women). Descriptive statistics about the final sample are reported in Table 1.1. Table 1.2 recapitulates the steps taken in the sample construction.

1.3.3 Employment statuses in INPS data

For each retiree in our sample, I can reconstruct reliable career trajectories since the entry in the labour market until retirement based on their *Estratto Conto*, i.e. their contributory history at INPS. Spells that fall outside the scope of *Estratti Conto* pertain to five main cases: (i) transition into temporary unemployment not covered by any allowance, (ii) transition into the informal labour market, (iii) transition into permanent

[6]In measuring the share of career spent contributing to the FPLD scheme, I exclude employment spells after retirement.

[7]The imbalance in observations across sexes is consistent with women's traditionally low participation to the Italian labour market.

unemployment/inactivity,[8] (iv) transition into the public sector / high-profile freelance jobs, (v) migration abroad. Since the focus of our analysis is on late-career trajectories, i.e. the last ten years prior to retirement, case (iv) can be ruled out since the likelihood of getting a public sector job (which in most of the cases implies passing a public examination) or turning into highly qualified professional activity in late adulthood is likely negligible (Contini & Leombruni 2006). While the data do not allow to directly address the case of migration abroad, statistics on the characteristics of Italian emigrants suggest that migration in later working life is also infrequent.[9] Case (i) can be properly identified if the individual disappears from the data and re-appears during a given year, of if s/he disappears and reappears during the following year. Unfortunately, the data do not allow to disentangle the remaining two cases. Long-term disappearance from the scope of *Estratti Conto* in the last years prior to retirement may be well due to transition into inactivity or into the informal labour market. I thus consider these two cases together as (formal) unemployment. With these caveats in mind, I identify 7 mutually exclusive employment statuses in which individuals may predominantly fall during a given year: (i) full-time dependent work, (ii) part-time dependent work, (iii) self-employment work[10], (iv) temporary suspension from work covered by wage subsidy public schemes (the so-called *Cassa Integrazione Guadagni*, or CIG), (v) unemployment (i.e. having worked less than 26 weeks during a year) covered, at least in part, by unemployment benefits, (vi) unemployment (i.e. having worked less than 26 weeks during a year) with no unemploy-

[8]It is important to stress that transition into permanent inactivity may be associated to so-called 'voluntary exodus', i.e. voluntary redundancy regulated by private agreements between older workers and companies by which the former decide to leave their job in exchange for support for the years leading to the retirement age. Voluntary exodus has become a policy issue in Italy after the 2011 pension reform, which raised the retirement age by 3-4 years abruptly, leaving about 300,000 workers which had opted for voluntary exodus with no job and no pension (Checchi & Leonardi 2015).

[9]According to the 2019 report on registration and de-regristration trends of the Italian resident population issued by the Italian National Institute of Statistics (ISTAT), the average age of Italian emigrants in 2018 was 33 in the case of men and 30 in the case of women. The share of Italian emigrants aged 50+ was 13%. See `https://www.istat.it/it/files//2020/05/Migrazioni_EN.pdf`

[10]A particular type of workers' category covered by INPS data is that of 'parasubordinate' workers, i.e. workers who fall formally under non-subordinate contractual arrangements while working, *de facto*, as dependent employees with reduced access to social welfare benefits. We classify spells of parasubordinate work as self-employment.

ment benefits,[11] and (vii) a residual category of statuses, as resulting from the *Estratti conto*, which do not fall into any of the former six (e.g. family leaves).

1.4 Methods

1.4.1 Sequence and cluster analysis

To identify late-career trajectories, I apply sequence analysis. According to this theoretical and methodological framework, individual trajectories can be represented as sequences of categorical events, or states (Abbott 1995). The first step in sequence analysis is generally to produce a matrix of dissimilarity between individual sequences, which can then be used as input for data-reduction techniques, such as cluster analysis. Groups of individual sequences identified through data reduction techniques can further be used as determinants, or consequences, of life course trajectories (Barban & Billari 2012).

I construct individual sequences of employment statuses for each individual in the final sample focusing on the 10 years prior to retirement. Specifically, I code each individual yearly-based observation according to the prevalent employment status, among the 7 outlined above. For instance, the trajectories of two individuals X and Y who transition, respectively, from full-time employment (FT) to unemployment without social allowances (UWB) in the last three years prior to retirement and from full-time employment to part-time employment (PT) in the last five years prior to retirement would look as follows:

[11] It should be noted that employment state vi) includes both individuals who spend >26 weeks overall in formal unemployment without unemployment benefits but worked for the rest of the year, and individuals who spend the entire year in format unemployment (i.e. do not appear in the dataset), in any of the ten years prior to retirement. While closely related, these two cases are conceptually different as the former category of individuals experience some transition in and out of unemployment during the year, while the latter do not. For this reason, in a preliminary version of the paper, I differentiated between the two, obtaining results for both cluster and survival analysis which were very similar to those presented in the current version of the paper. For the sake of simplicity, I thus consider the two cases as falling into the same employment state.

$$X : (\text{FT, FT, FT, FT, FT, FT, FT, UWB, UWB, UWB})$$

$$Y : (\text{FT, FT, FT, FT, FT, PT, PT, PT, PT, PT})$$

If an individual experiences multiple employment statuses during a year, she is assigned to the status which lasted the longest in terms of weeks. I use the `R` package `TraMineR` developed by Gabadinho et al. (2011) to analyze the individual sequences so constructed. Given that employment trajectories of men and women tend to differ, I perform the analysis separately for each gender.

I first compute pairwise dissimilarity matrices for men and women using the dynamic Hamming algorithm to measure distance between individual sequences. The dynamic Hamming method has been proposed as an alternative to traditional optimal matching techniques. In optimal matching, the dissimilarity between two sequences is measured by considering how much effort is needed for transforming one sequence into the other (Barban & Billari 2012). Transformation entails three basic operations (insertion, deletion and substitution) to which specific costs should be assigned. Critics of optimal matching generally point at the difficulty of providing solid theoretical background to the determination of these costs (Gauthier et al. 2009). They also stress failure of optimal matching to account for non-linear dependency over time. That is, by relying on transformation costs which are the same at any point of the sequence, and independent of the direction, optimal matching ignores the ordering of sequences. This implies, for instance, that optimal matching treats transitioning from employment to unemployment as equivalent to transitioning from unemployment to employment, and transitioning from employment to unemployment as equally costly regardless of whether it occurs at age 50 or 65. The dynamic Hamming method proposed by Lesnard (2006) addresses both shortcomings. First, it does not use insertions or deletions (for this reason, it can exclusively handle sequences of equal length). Second, it employs substitution costs which depend on position t in the sequence and which are derived from transition rates between possible states observed in the sample at successive positions. Following (Lesnard 2006), the dynamic Hamming time-dependent substitution cost between states a and b at position t is defined as:

$$s_t(a,b) = 4 - p_t(b|a) - p_t(a|b) - p_{t+1}(b|a) - p_{t+1}(a|b) \tag{1.1}$$

where $p_t(b|a)$ is the probability of transitioning from a to b between $t-1$ and t, and $s \in [0,4]$. The higher (lower) the transition rate between states a and b between $t-1$ and t, and between t and $t+1$, the lower (higher) the substitution cost between a and b at t. Given its timing sensitivity, the dynamic Hamming method is particularly useful for applications in which the exact timing/position of states within sequences is theoretically important, as in the case of retirement trajectories (Aisenmbrey & Fasang 2010).

I use the dissimilarity matrices so derived to identify ideal-type late-career trajectories for men and women by means of hierarchical cluster analysis. As linkage criterion, I rely on the Ward's method, which minimizes the within-cluster variance. I determine the number of clusters based on three criteria: (i) the observation of theoretically meaningful clusters, (ii) saturation (i.e. whether the addition of a new cluster is just another version of those already existing), and (iii) sufficient number of observations in each cluster.[12] I further evaluate the adequacy of clusters so determined by relying on the average silhouette width (ASW) criterion, weighted by the number of sequences in each cluster, which measures the coherence of assignment of each sequence to a given cluster. Average silhouette width ranges between 1 and -1, with 1 indicating very good clustering and values >0 acceptable quality of clustering (Torssander & Almquist 2017). Based on these criteria, as further detailed in Section 1.5, I eventually end up with a six-cluster solution for

[12] As a rule of thumb, I considered 1000 individuals to make a sufficient number.

men and a seven-cluster solution for women. In both cases, the average silhouette width (weighted) is close to 0.6, which indicates good quality of clustering. Plots of average silhouette width (weighted), for different cluster solutions, are reported in Figure A.1 in the Appendix, for men and women respectively. Using partitioning around medoids (PAM) criterion (Kaufman & Rousseeuw 2005) as alternative clustering technique yields analogous clustering solutions, in terms of qualitative differences between clusters, cluster assignment and consistency within clusters (Figure A.2 and Figure A.3 in Appendix A).

1.4.2 Survival analysis

I examine the relationship between late-career employment trajectories and post-retirement mortality by using the identified clusters as explanatory variables in survival analysis, where the focal event is death. Individuals are followed from the year of retirement until 2018 or until the year of death if this occurs before 2018. My outcome of interest is a dichotomous variable taking value 1 if the individual is dead by the end of the year, 0 if she is still alive. Since the data do not allow to identify the precise date of death, but only whether any individual is dead/alive by the end of each year, I turn to to discrete-time survival analysis.[13] Specifically, I opt for the complementary log-log model that best fits the case of survival times which are continuous in nature but come grouped or banded into intervals (Jenkins 2005). The model specification for the interval hazard rate, for each sex, looks as follows:

$$log(-log[1 - h_j(X)]) = \beta' X + \gamma_j \tag{1.2}$$

or

[13]Information about the year of death, when applicable, come from the *Anagrafica* archival dataset. One may alternatively derive more precise information about the timing of death (month and year) by looking at the flows of pension instalments. Still, time mismatches between death and interruption of the pension flow can always occur. Thus, I take a conservative approach and use death information provided directly by INPS.

$$[h_j(X) = 1 - exp[-exp(\beta' X + \gamma_j)] \tag{1.3}$$

where $h(j, X)$ defines the hazard rate of dying over the interval j, X are the covariates with the related coefficients β, and γ_j are parameters that summarize the duration dependence in the interval hazard. In this case, I opt for a fully non-parametric baseline hazard. This implies creating duration-specific-interval dummy variables, one for each spell year at risk. The main explanatory variables are the dummies for the various ideal-type late-career trajectories identified through cluster analysis. Baseline controls include year of birth to account for cohort effects, region of residence, and a set of dummies for the prevalent type of occupational position held throughout the observable career (blue-collar, white-collar and managers[14]). Indeed, both geography and occupational history are likely to influence late-career individual trajectories, just as they happen to be well-established drivers of health inequalities (Leombruni et al. 2010; Petrelli et al. 2019). Baseline controls include also: age of appearance in the *Estratti Conto* records, which can be interpreted as a proxy of age of entry into the labour market, age at retirement, and a dummy for whether the individual records any employment spell after retirement.[15] The main threat to the validity of my analysis comes from unobservable health conditions which can affect both retirement trajectories and post-retirement mortality. I seek to address this reverse causality concern by controlling for the reception of pension benefits associated to physical or mental impairment, occurring prior to the attainment of old-age/seniority pension. I do so by including a set of dummies for the reception of (a) disability pension benefits, (b) indemnity pension benefits and (c) social pension benefits (see Appendix A.1 for details about pension benefits). I also include a dummy for the reception, prior to retirement, of survivors' pension benefits. While not directly related to recipient's

[14] This category include both '*quadri*', who hold middle-management positions, and '*dirigenti*', who hold executive positions.

[15] In Italy, old-age/seniority pension benefits are fully compatible with earnings from autonomous or dependent work since January 2009. Prior to this date, paid work after retirement implied a reduction in pension benefits.

physical or mental impairment, survivors' benefits track the occurrence of a major event, i.e. the loss of one's spouse, which may affect both health conditions (Roelfs et al. 2012) and labour supply decisions (Fadlon & Nielsen 2020; Giupponi 2019).

1.5 Results

1.5.1 Sequence analysis

Figure 1.1 reports the transversal state distribution, i.e. the percentage of individuals in each employment state, in the ten years prior to retirement for men and women who retired between 2001 and 2018. In each figure, the horizontal axis measures the time to retirement expressed in terms of years. Full-time employment is the most frequent employment status both for older male and female workers, although its relative importance decreases as retirement approaches, notably in the case of women. The most noticeable difference across sexes is the larger incidence of part-time work and long-term unemployment/inactivity prior to retirement among women as compared to men.[16] This is line with the weaker attachment of Italian women to the labour market due, *inter alia*, to gender asymmetries in caring responsibilities in a context traditionally marked by 'familistic' welfare (Esping-Andersen 1990). The limited relevance of self-employment is not surprising given that my sample consists of people who paid contributions as private sector employees for most of their career.

As noted above, hierarchical cluster analysis delivers a six-cluster solution for men and a seven-cluster solution for women. Figure 1.2 displays clusters of individual sequences, ordered based on the distance from the most frequent sequence in each cluster, for men and women respectively.[17] As one can notice, the composition of clusters across sexes

[16] It should be noted that INPS data do not allow to measure the proportion of part-time work compared to full-time work, nor they allow to qualify part-time work as a voluntary or involuntary choice. This limits both the room for gender comparison and for analysing the influence of part-time work trajectories on subsequent health outcomes.

[17] For visualization reasons, 500 representative sequences (randomly drawn) are shown for each cluster. Graphs reporting all sequences in each cluster are available upon request.

is qualitatively similar. Cluster "Full-time employment" is clearly dominated by stable late-career employment trajectories, corresponding to the paradigm of smooth transition from full-time employment directly into retirement. At the national level the majority of men falls into this cluster (60%), while this is the case for about 39% of women. Still, as we further illustrate below, sizeable heterogeneities exist across macro-regions (Table 1.3). Cluster "Self-employment" features sequences characterized mostly by spells of self-employment work. This is the case of individuals (about 7% and 6% of sampled male and female retirees) who used to hold a dependent job, and who spent the ten years prior to retirement working mostly as self-employed. Cluster "Full-time employment/unemployment without benefits" is characterized by sequences featuring early stable employment which then gives way to spells of unemployment not covered by social benefits prior to retirement. About 6.7% of male retirees and 5.5% of female retirees in our sample fall into this case. Cluster "Full-time employment/unemployment with benefits" features sequences where full-time work gives way to unemployment spells, mostly covered by unemployment benefits. This cluster concerns about 10.7% of men and 8.1% of women in our sample. Cluster "Unemployment without benefits" is characterized by trajectories dominated by spells of (formal) unemployment not covered by social benefits along the entire decade preceding retirement. This trajectory fits the case of individuals who may be unable to find a (formal) job after dismissal or who may voluntarily drop out of the labour market. About 14% of male retirees and 26.6% of female retirees in our sample fall into this case. Cluster "Part-time employment" is characterised by the prevalence of part-time work spells. The share of men fall into this category is minimal (0.9%), while it is sizeable in the case of women (13.1%). Last, cluster "Full-time/part-time employment", the only women-specific cluster, is characterized by trajectories where full-time work gives way to part-time work prior to retirement. About 2.8% of female retirees in our sample follow this late-career trajectory.

Regional disparities in the distribution of ideal-type late-career trajectories are markedly pronounced (Table 1.3). In the case of men, slightly less than 70% of male retirees in the North of the peninsula, traditionally richer and economically more dynamic than the South (Felice 2018), fall in the full-time dependent work career path, while this is the case for less than 50% of the retirees in the Southern regions. Likewise, male pensioners from the Centre-North have remarkably less chance of going through spells of unemployment compared to the male retirees living in the South. Considerable regional disparities exist also in the case of women. The proportion of women belonging to cluster "Full-time employment" in the North-West is more than 25pp higher compared to women in the South. The same holds true for the "Part-time employment" trajectory which is three times more common in the North than in the South. By contrast, women in the South are three to five times more likely than Northern ones to go predominantly through spells of unemployment/inactivity, not covered by any social allowance. In fact, the proportion of women spending most of the ten years prior to retirement out of the formal labour market is large in all macro-regions, compared to men, but particularly so in the South. Still, it is worth remembering that our data do not allow to distinguish between long-term unemployment, voluntary inactivity and transition into the informal labour market. It is possible that some of the Southern older workers who result as formally unemployed are actually active in the shadow economy.

There are also some noticeable trends in the prevalence of ideal-type career trajectories over the selected years (2001-2018). As documented by Figure 1.3, there is an increase in the share of women with retirement trajectories marked by full-time and part-time employment, and a sizeable decline of those spending most of the decade prior to retirement into formal unemployment. These trends can be associated to pension reforms which raised the statutory retirement age, making increasingly costly for women to leave the labour market while waiting to meet the age requirements to claim pension benefits.[18] In the case of men, we observe an increase in the share of men who go through

[18]These trends become particularly visible after the *Fornero* reform (Law-decree n. 201/2011), which provided for a sharp rise in the statutory retirement age, particularly in the case of women, starting from 2012 (Moscarola et al. 2016).

spells of unemployment covered by unemployment benefits, notably after the 2011 debt crisis, to the detriment of direct transition from full-time work to retirement. Such trend could be related to policy measures introduced in 2012 and 2015 which expanded unemployment benefit provisions for private sector-employees, and strongly relaxed access requirements.[19] Albeit quantitatively marginal, we observe also an increase in the share of men with retirement trajectories marked by part-time employment. While some of the observed trends may stem from sluggish economic conditions in the 2010s (Lorenti et al. 2019), they are also likely to reflect structural changes in the Italian labour market induced by reforms aimed at increasing the length of working life (Carta & Dephilippis 2021).

Overall, sequence and cluster analysis confirm that a non-negligible share of individuals who retired in Italy between 2001 and 2018 experienced some form of employment instability, measured in terms of deviation from the full-time employment paradigm. Also, although the data do not allow for a proper distinction between unemployment and employment in the informal labour market, a strong South-North divide emerges from the analysis, deviations from full-time employment in later working life being markedly more frequent in Southern than in Northern regions.

1.5.2 Survival analysis

The outcome of interest is a dichotomous variable taking value 0 if the individual is still alive at the end of the year, 1 if she is dead. As noted above, I do not impose any constraint on the baseline hazard. That is, in each regression I include as many dummies as the maximum survival time observed. Since the earliest year of retirement in the dataset is 2001, and there are individuals from the 2001 retirement cohort who are still alive by the

[19] More specifically, the 2012 reform, also known as the Fornero unemployment benefit reform, increased wage replacement rates for first six months of unemployment from 60% to 75%. Three years later, the 2015 reform markedly eased requirements for requiring access to unemployment benefits, setting a minimum of (i) 13 weeks of social security contributions in the four years prior to the request, and (ii) 18 days of effective work in the year prior to the request (the former regime imposed a minimum of 12 months of contributions in the two years prior to the request).

end 2018, the maximum survival time is 18 years. In all tables, estimated coefficients are reported in the exponentiated form, in order to make them readily interpretable in terms of hazard ratios. Parentheses report robust standard errors.

Table 1.4 displays results from complementary log-log regressions for men. The main explanatory variables of interest are the dummies for the ideal-type late-career employment trajectories. The cluster of reference is always the ideal-type trajectory characterized by full-time employment. For reasons of space, the estimated hazard ratios for the duration-specific interval dummies are not shown.[20] Column 1 reports results when controlling for year of birth only. In this baseline model, all retirement trajectories deviating from full-time private employment come with higher post-retirement mortality risk (hazard ratios > 1). In Column 2, I add controls for a number of potential confounders. First of all, I control for macro-region of residence, as geography in Italy is an important factor influencing both employment perspectives (Aimone Gigio et al. 2021) and health outcomes (Petrelli et al. 2019). Second, I control for prevalent type of occupational position held throughout the observable career as the latter can plausibly influence late-career trajectories and is a well-established determinant of mortality (Leombruni et al. 2010). I further control for age at first job, which can be seen as a proxy for educational level, that is an important factor influencing both senior workers' employment prospects (Visser et al. 2016) and health outcomes (Mackenbach et al. 2019). I also control for age at retirement and post-retirement employment which may reflect factors likely related to both late-career employment trajectories and post-retirement mortality. Age at retirement may reflect both unobserved health status and life-long employment history. For instance, retiring at a relatively early age may stem from poor health, but also from an uninterrupted career which gives access to pension benefits before reaching statutory retirement age based on years of contributions. Similarly, post-retirement employment may reflect (good) unobserved health (Wahrendorf et al. 2017), but also poor economic conditions

[20] Overall, the magnitude of the estimated hazard ratios of duration-specific-interval dummies suggests that the hazard increases as time from retirement elapses. Full results, including estimated baseline hazards, are reported in Table A.1 and Table A.2 in the Appendix.

(Principi et al. 2012).

Controlling for all these potential confounders make the hazard ratios decrease in magnitude, but all late-career trajectories deviating from full-time employment remain positively associated to higher post-retirement mortality risk. The greatest risk (+27% with respect to the cluster of reference) is attached to the "Full-time/Unemployment without benefits" trajectory, suggesting that the scarring effect of unemployment is particularly harmful if it follows full-time employment and is not covered by unemployment allowances. Dwelling shortly on control variables, a few results are worth stressing. First, higher age at first job comes with lower post-retirement mortality risk, which is consistent with the interpretation of it being a proxy for educational level. Second, geographical disparities in post-retirement survival are quite pronounced. Male retirees living in the North-East have a higher (+7%) post-retirement mortality risk compared to male retirees living in the Centre, while the opposite holds for those living in the South (-12%) and in the Islands (-16%). This inverse North-South post-retirement mortality gradient is in line with findings from previous works investigating mortality patterns among Italian pensioners formerly employed in the private sector on the basis of INPS archival data (Belloni et al. 2012; Lallo & Raitano 2018; Leombruni et al. 2010).[21] Third, having a non-blue-collar background has a strongly protective effect in terms of lower post-retirement mortality risk (-25% and -19% for former managers and white-collar employees respectively, as compared to former blue-collar workers), in line with previous works (Leombruni

[21] It is worth stressing that analyses based on linked census and mortality records find evidence of higher mortality risk in the general population in southern Italy, compared to the rest of the country (Petrelli et al. 2019). A plausible explanation for the inverse North-South mortality gradient documented on the basis of INPS archival data lies in the composition of INPS population which consists, as noted, of private sector workers. As such, the INPS population is not fully representative of the Italian general population, since it excludes individuals who work in the public sector, those who work in the black market, and those who have never joined/have dropped out of the labour force. I advance two possible reasons for the inverse North-South mortality gradient among (former) private sector workers. First, it is possible that individuals with poor/vulnerable health are less likely to join the private sector in the Centre-South as compared to the North, maybe because they are more likely to join the public sector or the black market, or to drop out of the labour force *tout court*. In other words, it is possible that private sector workers in the Centre-South are selectively healthier as compared to the North. Second, it is possible that private sector workers living in the North are exposed to risk factors, such as employment in specific industries, which make them more vulnerable compared to their counterparts in the Centre-South.

et al. 2010).

The main threat to the validity of these results is represented by reverse causality dynamics whereby unobservable health conditions affect both retirement trajectories and post-retirement mortality. It is possible, indeed, that individuals who deviate from the full-time work trajectory do so for health reasons, implying an upward bias in the hazard ratios for the various clusters. This is also suggested by the higher incidence of recipients of allowances payable upon the occurrence of physical/mental impairment among individuals who deviate from the "Full-time employment" trajectory (Table A.3 in the Appendix). I seek to address this issue by controlling for: (i) the reception of allowances related to physical/mental impairment prior to retirement, (ii) the total number of sickness/injury leave episodes in the ten years prior to retirement, (iii) the cumulative number of full weeks in sickness/injury leave in the ten years prior to retirement.[22] Results, shown in Column 3, suggest that some reverse causality is indeed at play, as witnessed by the loss of magnitude and statistical significance of our hazard ratios of interests and by the strong positive association between the various types of allowances and post-retirement mortality risk. Still, all retirement trajectories, but the one dominated by part-time employment, remain significantly associated to higher post-retirement mortality risk. In particular, men who spend most of the decade prior to retirement in formal unemployment have a post-retirement mortality risk which is 13% higher than those who transition from full-time private employment into retirement. Going from full-time employment through unemployment is associated to an increase in post-retirement mortality risk by 9.9%, if unemployment is covered by allowances, and by 8.0%, if it is not. This difference could is surprising as one may expect unemployment allowances to have, *ceteris paribus*, a protective effect on health. However, the difference between the two estimated coefficients is statistically non-significant (p> $\chi^2 = 0.66$)

Table 1.5 reports results from complementary log-log regressions for women. Column 1

[22] It is important to stress that in the data, sickness/injury leave episodes lasting less than one week are reported as lasting zero weeks.

shows results for the baseline model where, in addition to the explanatory variables of interests, I control for the year of birth only. In this baseline specification, retirement trajectories featuring unemployment spells not covered by allowances are strongly associated to higher post-retirement mortality risk. A positive association emerges also in the case of trajectories dominated by self-employment and of trajectories characterized by transition from full-time to part-time private employment. When adding controls for macro-region of residence, prevalent type of occupation, age at first job, age at retirement, and post-retirement employment, also the trajectory characterized by the transition from full-time private employment to unemployment covered by allowances comes with a higher post-mortality risk compared to the cluster of reference. It is worth noting that the relationship between post-retirement mortality and control variables differs between men and women. Indeed, in the case of women, higher age at first job results in a higher post-retirement mortality risk. To the extent that age at first job is a reliable proxy of educational attainment, this result is consistent with previous studies finding a non-linear, when not inverse, gradient between education and health for Italian women (Leombruni et al. 2010). This interpretation is corroborated by the fact that female retirees with former white-collar and managerial jobs do not display a significantly lower risk than those with a blue-collar background. The relevance of macro-region of residence is attenuated, too. To address reverse causality concerns, in Column 3, I include controls for the reception of allowances related to physical or mental impairment prior to retirement, the total number of sickness/injury leave episodes in the ten years prior to retirement, and the cumulative number of full weeks in sickness/injury leave in the ten years prior to retirement. Unlike men, the retirement trajectory which comes with the strongest post-retirement mortality risk is the one dominated by self-employment spells (+21.0% compared to the "Full-time employment" cluster), followed by the trajectory characterized transition from full-time private employment to unemployment covered by allowances (+18.5%) and by the trajectory marked by prolonged unemployment (+12.2%). Being employed part-time or transitioning from full-time employment into unemployment not covered by allowances has no appreciable association with post-retirement mortality risk. It is possible that women who leave full-time employment entering formal unemployment do so voluntarily, possibly upon own partner's retirement (Bloemen et al. 2019; Pozzebon & Mitchell 1989), even if they have not reached the requirements for accessing pension benefits, and this may be actually beneficial for their health (Zang 2020).

In Table 1.6, I allow for heterogeneity in the association between ideal-type trajectories and post-retirement mortality based on prevalent occupational category and macro-region of residence. Columns 1 to 2 report results for men. In Column 1, clusters are interacted with dummy variables taking value 1 if the prevalent occupational category throughout the observable category is white-collar or manager. Results suggest that the association between retirement trajectories and post-retirement mortality risk does not significantly differ across occupational groups, broadly defined. In Column 2, I interact clusters with a dummy taking value 1 if the macro-region of residence is either North-East or North-West. I find that the positive association between post-retirement mortality risk and trajectories marked by unemployment spells, not covered by allowances, is stronger for men living in the North, where deviations from the full-time private employment for mature workers is less frequent, as compared to men living in the rest of Italy (Centre, South and Islands). In the case of women, the relationship between retirement trajectories and post-retirement mortality differs along both the occupational and geographical dimension. Column 3 in Table 1.6 shows that holding a white-collar background (but not a managerial one) mitigates the adverse consequences of going through trajectories deviating from full-time employment as compared to holding a blue-collar background. For instance, white-collar women who go through prolonged unemployment prior to retirement have a 15.5% (1.228*1.189*0.791) higher mortality risk compared to blue-collar women in full-time employment (i.e. the reference group). In contrast, blue-collar women who go through prolonged unemployment prior to retirement have a 22.8% higher mortality risk compared to blue-collar women in full-time employment. Likewise, white-collar and blue-collar women who transition from full-time employment to unemployment not covered by allowances prior to retirement have a post-retirement mortality risk which is, respectively, 7.4% (1.143*1.189*0.689) lower and 14.3% higher compared to women with a blue-collar background in full-time employment. A possible interpretation of these findings is that while for white-collar women leaving full-time employment prior to retirement could be a personal choice, for women with a blue-collar background it could be the result of involuntary circumstances. Finally, Column 5 documents that also in the case of women prolonged unemployment prior to retirement scars more in the North, where this trajectory is less common compared to other areas of the country.

1.6 Sensitivity analyses

I corroborate results from survival analysis through a series of robustness checks. Column 1 of Table 1.7 and Table 1.8 reports results from complementary log-log regression where I include a set of dummies for cohort- and sex-specific quintiles of average inflation-adjusted gross income in the ten years prior to retirement, observable in the *Estratti Conto*, in addition to the full set of baseline controls, for men and women respectively. Although average gross income in the decade prior to retirement might be endogenous to the type of late-career employment trajectory, it serves as a proxy for life-time income and socio-economic status, which may plausibly correlate with both retirement trajectories and health. Results for both men and women are qualitatively and quantitatively robust to the inclusion of these new variables. It is worth stressing that in the case of men there is an inverse, albeit not perfectly linear, relationship between income quintile and mortality. This relationship does not hold, instead, in the case of women. In fact, women with higher income are exposed to higher mortality risk than women at the bottom of the income distribution, consistently with evidence of inverse socio-economic gradient reported in previous studies focusing on Italian women (Costa et al. 2017).[23] It should also be noted, though, that pre-retirement average gross income could also be interpreted as a mediating factor. Under this interpretation, results suggest that employment instability relates to post-retirement mortality through channels other than income dynamics. Columns 2 to 6 of Table 1.7 and Table 1.8 report results from specifications where, in addition to full baseline controls, I control also for the total number of transitions across employment states in the decade prior to retirement, the number of transitions to unemployment with and without unemployment allowances, and the number of yearly spells spent into unemployment, with and without unemployment allowances. The purpose of these additional specifications is to assess whether retirement trajectories matter on top of the simple characterization of single (un)employment events. Overall, results for both men and women are robust to these checks. The only exception is that of the retirement trajectory characterized by transition from full-time employment to unemployment covered by unemployment allowances: when controlling for the number of spells spent into unemployment covered by unemployment allowances, its association with post-retirement male mortality risk loses statistical significance (Table 1.7, Column 5). However, it is worth stressing that these results should be interpreted cautiously as these controls are part of the definition of late-career employment trajectories themselves.

Table 1.9 displays the results of a model where I interact clusters for ideal-type late-career trajectories with a dummy taking value 1 if the individual receives pension allowances related to physical or mental health impairment.[24] The purpose of this model is to further examine the interplay between health selection and employment trajectories. Specifically, this model allows to test whether individuals who deviate from full-time employment trajectories are exposed to higher post-retirement mortality rate even if they do not receive any allowance related to physical or mental health impairment, i.e. they are presumably in good health, upon retirement. Results suggest that, indeed, healthy individuals who

[23] Another explanation could be that women's own income is not a good proxy of socio-economic status, which might be better measured through household overall income.

[24] As explained in Appendix A.1, such pension allowances include disability pension benefits, indemnity pension benefits and social pension benefits.

go through trajectories marked by deviations from full-time employment face higher post-retirement mortality compared to healthy individuals in the full-time employment cluster, in the case of both men and women. Results suggest also that health issues, proxied by the reception of health-related pension allowances, magnify the post-retirement mortality risk of going through trajectories deviating from full-time employment in the case of men, but not in the case of women.

Finally, I carry out sequence and cluster analysis using semesters (six-months periods), rather than years, as time unit for identifying employment trajectories. These additional analyses are meant to alleviate concerns that coding yearly spells with the prevalent employment status over any given year may mask important heterogeneities in late-career employment patterns. While the semester-based approach is still ignoring some heterogeneity, since individuals experiencing multiple employment states over a given semester are assigned to the state lasting longer in that semester, such issue should be less serious than in the year-based approach. I compare the semester-based approach with the baseline one by looking at the adequacy of different cluster solutions, clusters' composition and cluster assignment, as well as the results from survival analysis where semester-based clusters are the predictors of interest. Overall, as further documented in Appendix A.4, the semester- and year-based approach yield similar results.

1.7 Discussion

The relationship between employment and retirement in advanced economies has deeply changed in the last decades. The old life-course paradigm of smooth transition from paid work to pension income has given way to late-career trajectories which have become differentiated and de-standardized. Scholars have started to investigate the socio-economic consequences of these changing patterns, whereby senior workers experience different employment conditions, along heterogeneous paths. Still, while there is already evidence that heterogeneity of retirement patterns may have important implications on income inequality of older people, little attention has been paid to the implications that these very same patterns may have on health inequality, and in particular on mortality and survival chances to old age. This work, examining the relationship between late-career employment trajectories and post-retirement mortality for a large sample of Italian retirees formerly employed in the private sector, is a first attempt to address this gap.

The main findings suggest that late-career trajectories marked by periods of unemployment, especially if prolonged and not covered by social allowances, are related to lower post-retirement survival chances, such relationship being mildly stronger for men as compared to women. Heterogeneity analysis further suggests that the scarring effect of prolonged unemployment for senior workers in Italy are particularly accentuated for individuals belonging to disadvantaged occupational categories, at least in the case of women. For disadvantaged women, prolonged unemployment in later life could be particularly harmful as it may add to the consequences of weak labor market attachment over the life course, jeopardizing their ability to secure adequate retirement income in old age, which ultimately impinges on health and survival chances (Leombruni et al. 2010). Heterogeneity analysis suggests also that the scarring effect of prolonged unemployment in later life is particularly harmful in areas where deviations from full-time employment trajectories are less frequent. This result speaks to studies bringing evidence on unemployment having greater detrimental effect on well-being the less there is of it around (Clark 2003). Previous works suggest that, in general, mechanisms through which employment instability may channel into higher post-retirement mortality risk could be multiple (Benach et al. 2015), including exposure to unhealthy working conditions over the life course (Quinlan et al. 2001), psycho-social stress (Muntaner et al. 2010), and material deprivation (Siegrist & Theorell 2006). Lack of information in INPS data does not allow, though, to directly test the role of such alternative channels.

While referring to a specific empirical context, the results of this work could be of interest for countries facing the challenge of stretching the length of working life to alleviate the pressures of increased longevity on their social security systems. First of all, they highlight the importance of policies aimed at improving the employment opportunities of displaced senior workers, and at ensuring adequate welfare assistance in case of prolonged unemployment. They also point at the need for policy solutions providing alternatives other than stepping out of the labor market before meeting pension requirements to senior workers who may struggle with full-time employment commitments. More specifically, they call for life-course-oriented labor market policies whereby workloads can be calibrated to life-phase-specific skills, abilities and needs. Facilitating shorter working hours, for instance, may not only encourage people to keep on working, but may actually enable them to do so (Eurofound 2016). While part-time work has been historically less common in Italy than in other European countries, the incidence of late-career trajectories marked by part-time employment among Italian senior workers has been increasing over time, especially among women. Reassuringly, I find that, after adjusting for potential confounders, late-career trajectories characterized by part-time employment are not significantly associated to higher post-retirement mortality risk compared to trajectories in full-time employment. This aspect is particularly relevant for senior female workers who may opt for part-time work to deal with caring commitments in a context where caring responsibilities, particularly towards the elderly, continue to reflect a gendered division of labor (Saraceno 2018).

Some caveats and limitations apply. First of all, lack of information concerning relevant socio-economic characteristics, such as education or marital/family status, limits the room for controlling for potential confounders. Failure to control for the unobserved factors could be problematic to the extent they affect both employment trajectories and post-retirement mortality outcomes, as this may lead to biased coefficient estimates. In addition, although I try to attenuate reverse causality concerns by dropping individuals who experience prolonged sickness- or injury-related leaves in any of the ten years prior to retirement and by controlling for the reception of pension benefits related to mental and physical impairment, health selection dynamics might still be at work. In particular, it is possible that individuals experiencing trajectories deviating from full-time employment choose so because they are unable to take on full-time employment commitments for health reasons. If this were the case, such individuals would be selectively less healthy, facing worse survival prospects after retirement. More in general, the empirical strategy I adopt does not allow to rule out endogeneity concerns completely. For this reason, the results cannot and should not be read in terms of causality. Second, lack of detailed information about individual lifestyle habits, post-retirement health and financial conditions reduces the scope for uncovering mechanisms whereby specific employment trajectories channel into lower survival. It is possible, for instance, that trajectories deviating from full-time employment, particularly if they materialize in the form of prolonged unemployment, are more likely conducive to old-age poverty, social isolation, and unhealthy behaviors, which may translate into higher post-retirement mortality risk. Third, post-retirement survival analysis is carried out on a relatively limited timespan. For this reason, the analysis may not fully capture the extent to which employment instability relates to survival chances in the long-term. This is particularly relevant for gender comparison as women tend to live longer than men. As such, the consequences of employment instability for the former may materialize later on over the life course as compared to the latter. Fourth, the analysis is based on a sample which is representative of individuals formerly, and formally, employed in the private sector in Italy, which limits the generalizability of results to the Italian population as a whole.

Despite its limitations, this study presents also a number of strengths. First, it uses ideal-type late-career trajectories, identified through sequence and cluster analysis, rather than individual features determining such trajectories (such as the number and duration of unemployment spells), to explain differences in post-retirement mortality risk. In doing so, it tackles a limitation of most existing works which investigate the relationship between health and labor market events overlooking the dynamic nature of the latter. Second, it makes use of longitudinal register data to track employment patterns. Differently from survey data asking for retrospective information, register data are not affected by recall bias. As such, they ensure greater reliability and precision in the reconstruction of employment trajectories. Moreover, the data used in this work allow to exploit information, such as the reception of unemployment benefits and allowances alike, which add to the completeness of employment biographies, but which are rarely accounted for by studies of this type. Third, while most existing studies employ subjective measures of health and well-being, I adopt an objective measure of health, i.e. mortality, which is not exposed to self-reporting issues. Future research may seek to provide further insights by means of more fine-grained analyses. For instance, one may account also for transitions into lower paid or lower status jobs. Information about previous firm type (e.g. small-medium vs large-sized) or industry of employment could also be used to build more accurate measures of occupational history, and allow for better investigation of drivers and consequences of late-career employment trajectories.

Tables

Table 1.1: Descriptive statistics - INPS LoSai sample

	All				Men				Women			
	Mean	Sd	Min	Max	Mean	Sd	Min	Max	Mean	Sd	Min	Max
Year of birth	1946.7	4.7	1931.0	1966.0	1946.4	5.0	1931.0	1966.0	1947.2	4.2	1931.0	1965.0
Retirement age	59.5	3.7	51.0	70.0	59.7	4.1	51.0	70.0	59.2	2.7	51.0	70.0
Age at first job	17.9	3.5	12.0	30.0	18.0	3.4	12.0	30.0	17.8	3.6	12.0	30.0
Work after retirement (%)	19.3	39.5	0.0	100.0	22.7	41.9	0.0	100.0	13.0	33.7	0.0	100.0
Macro-region of residence (%)												
North-East	20.1	40.1	0.0	100.0	18.5	38.8	0.0	100.0	23.2	42.2	0.0	100.0
North-West	31.9	46.6	0.0	100.0	30.8	46.2	0.0	100.0	33.8	47.3	0.0	100.0
Centre	18.5	38.8	0.0	100.0	18.2	38.6	0.0	100.0	19.0	39.3	0.0	100.0
South	19.9	39.9	0.0	100.0	21.2	40.9	0.0	100.0	17.5	38.0	0.0	100.0
Islands	9.6	29.5	0.0	100.0	11.3	31.7	0.0	100.0	6.5	24.6	0.0	100.0
Prevalent occupation (%)												
Blue-collar	68.0	46.7	0.0	100.0	70.8	45.5	0.0	100.0	62.8	48.3	0.0	100.0
White-collar	30.0	45.8	0.0	100.0	26.7	44.2	0.0	100.0	36.2	48.0	0.0	100.0
Manager	2.0	14.1	0.0	100.0	2.6	15.8	0.0	100.0	1.0	10.1	0.0	100.0
Reception of pension benefits (%)												
Disability pension	4.2	20.1	0.0	100.0	4.5	20.6	0.0	100.0	3.8	19.2	0.0	100.0
Survivor pension	3.4	18.2	0.0	100.0	1.2	10.8	0.0	100.0	7.6	26.5	0.0	100.0
Indemnity pension	2.9	16.9	0.0	100.0	4.2	20.0	0.0	100.0	0.6	7.9	0.0	100.0
Social pension	1.3	11.4	0.0	100.0	1.5	12.3	0.0	100.0	0.9	9.5	0.0	100.0

Notes. Work after retirement measures whether the individual records any employment spell after retirement. North-East, North-West, Centre, South and Islands are dummies for the macro-region of residence as of 2018 or as of the year of death, if this occurs earlier. Blue-collar, white-collar and manager are dummies for the prevalent occupational status observed in the *Estratti Conto* dataset over entire working life. Disability pension, Survivor pension, Indemnity Pension and Social pension are dummies measuring whether individuals receives these types of pension benefits (see Appendix A.1) prior to retirement.

Table 1.2: Steps in sample construction

	Men	Women	Tot
Initial sample (unique individuals, N)	302,794	198,426	501,220
Excluded N with most of contributions not to FPLD	100,456	50,685	151,141
Excluded N with no contributory spells in 10 yrs before retirement	47,429	64,552	111,981
Excluded N who retire before 50 or after 70	2,485	768	3,253
Excluded N reporting sickness/injury in 10 years before retirement	936	921	1,857
Excluded N who appear in the *Estratti Conto* age < 12 or age > 30	4,510	3980	8,490
Final sample	146,978	77,520	224,498

Notes. FPLD (*Fondo Pensione Lavoratori Dipendenti*) is the INPS-managed pension scheme of private employees.

Table 1.3: Cluster distribution by macro-region and gender

Cluster	North-East	North-West	Centre	South	Islands	Total
Men						
Self-employment	2535 (9.1%)	3932 (8.7%)	1946 (7.4%)	1208 (3.9%)	822 (4.9%)	10443 (7.1%)
Full-time employment	19314 (69.6%)	30443 (67.2%)	16731 (63.5%)	14645 (47.7%)	7634 (45.2%)	88767 (60.4%)
Part-time employment	212 (0.8%)	331 (0.7%)	262 (1.0%)	415 (1.4%)	128 (0.8%)	1348 (0.9%)
Unemployment without benefits	2293 (8.3%)	3074 (6.8%)	2762 (10.5%)	7758 (25.3%)	4981 (29.5%)	20868 (14.2%)
Full-time empl./Unempl. without benefits	1401 (5.0%)	2373 (5.2%)	1889 (7.2%)	2731 (8.9%)	1451 (8.6%)	9845 (6.7%)
Full-time empl./Unempl. with benefits	1999 (7.2%)	5122 (11.3%)	2773 (10.5%)	3951 (12.9%)	1862 (11.0%)	15707 (10.7%)
Total	27754 (100.0%)	45275 (100.0%)	26363 (100.0%)	30708 (100.0%)	16878 (100.0%)	146978 (100.0%)
Women						
Self-employment	1271 (6.9%)	1961 (7.4%)	863 (5.8%)	369 (2.9%)	197 (4.1%)	4661 (6.0%)
Full-time employment	7791 (42.0%)	11987 (45.0%)	6476 (43.7%)	2223 (17.5%)	1438 (29.6%)	29915 (38.6%)
Part-time employment	3084 (16.6%)	3806 (14.3%)	1992 (13.5%)	639 (5.0%)	412 (8.5%)	9933 (12.8%)
Unemployment without benefits	3778 (20.4%)	4072 (15.3%)	2982 (20.1%)	7613 (60.0%)	2176 (44.8%)	20621 (26.6%)
Full-time empl./Unempl. without benefits	1015 (5.5%)	1494 (5.6%)	956 (6.5%)	551 (4.3%)	230 (4.7%)	4246 (5.5%)
Full-time empl./Unempl. with benefits	1067 (5.8%)	2611 (9.8%)	1116 (7.5%)	1176 (9.3%)	317 (6.5%)	6287 (8.1%)
Full-time/Part-time employment	540 (2.9%)	700 (2.6%)	418 (2.8%)	108 (0.9%)	91 (1.9%)	1857 (2.4%)
Total	18546 (100.0%)	26631 (100.0%)	14803 (100.0%)	12679 (100.0%)	4861 (100.0%)	77520 (100.0%)

Notes. The macro-region Islands include the Sicily and Sardinia regions.

Table 1.4: Post-retirement mortality and ideal-type late career trajectories
Men

Death	(1) Model 1	(2) Model 2	(3) Model 3
Cluster (Ref: Full-time employment)			
Self-employment	1.122***	1.086**	1.095**
	(0.034)	(0.034)	(0.034)
Part-time employment	1.546***	1.180†	1.155
	(0.137)	(0.105)	(0.103)
Unemployment w/o benefits	1.498***	1.246***	1.134***
	(0.032)	(0.028)	(0.027)
Full-time empl./Unempl. w/o benefits	1.502***	1.269***	1.080*
	(0.043)	(0.037)	(0.033)
Full-time empl./Unempl. with benefits	1.111***	1.088**	1.099**
	(0.034)	(0.034)	(0.034)
Year of birth	0.927***	0.979***	0.973***
	(0.002)	(0.002)	(0.002)
Age at first job		0.984***	0.985***
		(0.003)	(0.003)
Retirement age		1.101***	1.087***
		(0.003)	(0.004)
Work after retirement		0.592***	0.612***
		(0.013)	(0.014)
Occupational status (ref: Blue-collar)			
Manager		0.750***	0.789***
		(0.042)	(0.045)
White-collar		0.810***	0.845***
		(0.017)	(0.018)
Macro-region (ref: Centre)			
North-East		1.072**	1.083**
		(0.028)	(0.029)
North-West		0.997	1.011
		(0.024)	(0.024)
South		0.884***	0.864***
		(0.022)	(0.021)
Islands		0.844***	0.845***
		(0.025)	(0.025)
Invalidity pension			1.924***
			(0.053)
Survivor pension			1.309***
			(0.083)
Indemnity pension			1.074†
			(0.042)
Social pension			2.211***
			(0.098)
# of weeks in sickness/injury leave			0.998
			(0.002)
# of sickness/injury leave episodes			1.010†
			(0.005)
Observations	1,431,429	1,431,429	1,431,429
Deaths	16,458	16,458	16,458

Notes. Results from complementary log-log models. All models include 18 duration dummies (baseline hazard). Coefficients are expressed in the exponentiated form (hazard ratios). Robust standard error in parentheses.
*** p<0.001, ** p<0.01, * p<0.05, † p<0.1.

Table 1.5: Post-retirement mortality and ideal-type late career trajectories
Women

	(1) Model 1	(2) Model 2	(3) Model 3
Cluster (Ref: Full-time employment)			
Self-employment	1.176*	1.209**	1.210**
	(0.080)	(0.083)	(0.083)
Part-time employment	1.101	1.017	0.992
	(0.064)	(0.060)	(0.059)
Unemployment w/o benefits	1.288***	1.215***	1.122**
	(0.052)	(0.052)	(0.049)
Full-time/Part-time employment	1.225†	1.174	1.118
	(0.131)	(0.126)	(0.120)
Full-time empl./Unempl. w/o benefits	1.202**	1.127†	1.007
	(0.085)	(0.080)	(0.072)
Full-time empl./Unempl. with benefits	1.116	1.153*	1.185*
	(0.077)	(0.081)	(0.084)
Year of birth	0.966***	0.997	0.992
	(0.005)	(0.005)	(0.005)
Age at first job		1.010*	1.008†
		(0.005)	(0.005)
Retirement age		1.089***	1.082***
		(0.009)	(0.009)
Work after retirement		0.644***	0.668***
		(0.037)	(0.038)
Occupational status (Ref: Blue-collar)			
Manager		0.766	0.781
		(0.135)	(0.137)
White-collar		0.983	1.008
		(0.035)	(0.036)
Macroregion (Ref: Centre)			
North-East		1.033	1.056
		(0.052)	(0.054)
North-West		1.109*	1.141**
		(0.052)	(0.053)
South		0.958	0.915
		(0.053)	(0.053)
Islands		1.021	1.003
		(0.073)	(0.073)
Disability pension			2.665***
			(0.157)
Survivor pension			1.109†
			(0.066)
Indemnity pension			0.899
			(0.202)
Social pension			5.219***
			(0.437)
# of weeks in sickness/injury leave			0.994*
			(0.003)
# of sickness/injury leave episodes			1.008
			(0.008)
Observations	775,508	775,508	775,508
Deaths	3,921	3,921	3,921

Notes. Results from complementary log-log models. All models include 18 duration dummies (baseline hazard). Coefficients are expressed in the exponentiated form (hazard ratios). Robust standard error in parentheses. *** $p<0.001$, ** $p<0.01$, * $p<0.05$, † $p<0.1$.

Table 1.6: Heterogeneity analysis

	(1) Men	(2) Men	(3) Women	(4) Women
Self-employment	1.066†	1.097*	1.362***	1.160
	(0.037)	(0.052)	(0.111)	(0.144)
Part-time employment	1.167	1.088	1.108	0.939
	(0.112)	(0.132)	(0.081)	(0.100)
Unemployment w/o benefits	1.117***	1.051†	1.228***	1.038
	(0.028)	(0.028)	(0.065)	(0.063)
Full-time empl./Unempl. w/o benefits	1.060†	1.026	1.143	0.949
	(0.037)	(0.039)	(0.099)	(0.109)
Full-time empl./Unempl. with benefits	1.093*	1.117**	1.293**	1.238*
	(0.040)	(0.044)	(0.115)	(0.131)
Full-time empl./Part-time empl.			1.231	1.319
			(0.156)	(0.226)
White collar	0.824***		1.189**	
	(0.022)		(0.070)	
Manager	0.738***		0.905	
	(0.056)		(0.276)	
Self-employment x Manager	1.175		0.625	
	(0.221)		(0.369)	
Part-time employment x Manager	1.324		0.462	
	(1.340)		(0.484)	
Unemployment w/o benefits x Manager	1.284†		0.876	
	(0.181)		(0.368)	
Full-time empl./Unempl. w/o benefits x Manager	0.994		1.470	
	(0.199)		(0.867)	
Full-time empl./Unempl. with benefits x Manager	0.960		0.693	
	(0.314)		(0.727)	
Full-time empl./Part-time empl. x Manager			-	
			-	
Self-employment x White-collar	1.121		0.723*	
	(0.085)		(0.116)	
Part-time employment x White-collar	0.878		0.754*	
	(0.236)		(0.096)	
Unemployment w/o benefits x White-collar	1.048		0.791**	
	(0.058)		(0.067)	
Full-time empl./Unempl. w/o benefits x White-collar	1.083		0.689*	
	(0.076)		(0.109)	
Full-time empl./Unempl. with benefits x White-collar	1.022		0.816	
	(0.070)		(0.117)	
Full-time empl./Part-time empl. x White-collar			0.799	
			(0.192)	
North		1.094***		1.086
		(0.025)		(0.066)
Self-employment x North		1.013		1.069
		(0.061)		(0.158)
Part-time employment x North		1.145		1.084
		(0.204)		(0.138)
Unemployment w/o benefits x North		1.208***		1.130
		(0.052)		(0.093)
Full-time empl./Unempl. w/o benefits x North		1.129*		1.097
		(0.066)		(0.159)
Full-time empl./Unempl. with benefits x North		0.908		0.909
		(0.058)		(0.127)
Full-time empl./Part-time empl. x North				0.773
				(0.169)
Observations	1,431,429	1,431,429	775,508	775,508
Deaths	16,458	16,458	3,921	3,921

Notes. Results from complementary log-log models. Dependent variable: death occurrence (0,1). Coefficients are expressed in the exponentiated form (hazard ratios). Robust standard error in parentheses. All models control for year at birth, age at first job, retirement age, work after retirement, reception of disability, survivors', indemnity, and social pension benefits, # of full weeks in sickness/injury leave in ten years prior to retirement, # of sickness/injury leave episodes in ten years prior to retirement. All models include 18 duration dummies (baseline hazard). Columns 1 and 4 include controls for macro-region of residence. Columns 2 and 4 include controls for occupational status.
*** $p<0.001$, ** $p<0.01$, * $p<0.05$, † $p<0.1$.

Table 1.7: Robustness checks - Men

	(1) Model 1	(2) Model 2	(3) Model 3	(4) Model 4	(5) Model 5	(6) Model 6
Cluster (Ref: Full-time employment)						
Self-employment	1.078*	1.091**	1.096**	1.098**	1.096**	1.091**
	(0.034)	(0.034)	(0.034)	(0.035)	(0.034)	(0.034)
Part-time employment	1.125	1.155	1.155	1.154	1.153	1.147
	(0.101)	(0.103)	(0.103)	(0.103)	(0.103)	(0.103)
Unemployment w/o benefits	1.131***	1.130***	1.135***	1.155**	1.134***	1.133***
	(0.031)	(0.027)	(0.027)	(0.055)	(0.027)	(0.026)
Full-time empl./Unempl. w/o benefits	1.075*	1.067*	1.078*	1.089*	1.077*	1.067*
	(0.033)	(0.035)	(0.033)	(0.039)	(0.033)	(0.033)
Full-time empl./Unempl. With benefits	1.110***	1.099**	1.088*	1.099**	1.070	1.087**
	(0.034)	(0.034)	(0.039)	(0.034)	(0.046)	(0.034)
Pre-retirement income quintile (Ref: Bottom)						
2nd quintile	1.064*					
	(0.031)					
3rd quintile	1.074*					
	(0.032)					
4th quintile	0.965					
	(0.031)					
Top quintile	0.867***					
	(0.031)					
# trans. to Unempl. w/o benefits		1.014				
		(0.015)				
# trans. to Unempl. with benefits			1.012			
			(0.021)			
# spells in Unempl. w/o benefits				0.998		
				(0.006)		
# spells in Unempl. w/o benefits					1.008	
					(0.009)	
# total transitions						1.012†
						(0.007)
Observations	1,431,429	1,431,429	1,431,429	1,431,429	1,431,429	1,431,429
Deaths	16,458	16,458	16,458	16,458	16,458	16,458

Notes. Results from complementary log-log models. Dependent variable: death occurrence (0,1). Coefficients are expressed in the exponentiated form (hazard ratios). Robust standard error in parentheses. All models controls for year of birth, age at first job, age at retirement, post-retirement employment, prevalent occupational category, macro-region of residence, reception of disability, indemnity, social and survivors' pension benefits prior to retirement, # of full weeks in sickness/injury leave in ten years prior to retirement, # of sickness/injury leave episodes in ten years prior to retirement. All models include also 18 duration dummies (baseline hazard).
*** p<0.001, ** p<0.01, * p<0.05, † p<0.1.

Table 1.8: Robustness checks - Women

	(1) Model 1	(2) Model 2	(3) Model 3	(4) Model 4	(5) Model 5	(6) Model 6
Cluster (Ref: Full-time employment)						
Self-employment	1.200*	1.176*	1.211**	1.185*	1.211**	1.192*
	(0.085)	(0.083)	(0.084)	(0.087)	(0.084)	(0.083)
Part-time employment	1.001	0.986	0.992	0.987	0.991	0.983
	(0.063)	(0.058)	(0.059)	(0.059)	(0.059)	(0.058)
Unemployment w/o benefits	1.188**	1.097*	1.122**	1.038	1.121**	1.114*
	(0.063)	(0.050)	(0.049)	(0.104)	(0.049)	(0.049)
Full-time/Part-time employment	1.120	1.107	1.117	1.115	1.116	1.085
	(0.121)	(0.119)	(0.120)	(0.119)	(0.120)	(0.118)
Full-time empl./Unempl. w/o benefits	1.011	0.955	1.005	0.975	1.003	0.979
	(0.073)	(0.074)	(0.073)	(0.079)	(0.073)	(0.073)
Full-time empl./Unempl. with benefits	1.184*	1.177*	1.177*	1.184*	1.152	1.157*
	(0.084)	(0.083)	(0.095)	(0.083)	(0.106)	(0.084)
Pre-retirement income quintile (Ref: Bottom quintile)						
2nd quintile	$1.108^{\dagger}$					
	(0.061)					
3rd quintile= 3	1.163*					
	(0.072)					
4th quintile= 4	1.108					
	(0.070)					
Top quintile= 5	1.150*					
	(0.078)					
# trans. to Unempl. w/o benefits		$1.057^{\dagger}$				
		(0.033)				
# trans. to Unempl. with benefits			1.008			
			(0.044)			
# spells in Unempl. w/o benefits				1.010		
				(0.011)		
# spells in Unempl. with benefits					1.008	
					(0.017)	
# total transitions						1.022
						(0.015)
Observations	775,508	775,508	775,508	775,508	775,508	775,508
Deaths	3,921	3,921	3,921	3,921	3,921	3,921

Notes. Results from complementary log-log models. Dependent variable: death occurrence (0,1). Coefficients are expressed in the exponentiated form (hazard ratios). Robust standard error in parentheses.All models controls for year of birth, age at first job, age at retirement, post-retirement employment, prevalent occupational category, macro-region of residence, reception of disability, indemnity, social and survivors' pension benefits prior to retirement, # of full weeks in sickness/injury leave in ten years prior to retirement, # of sickness/injury leave episodes in ten years prior to retirement. All models include also 18 duration dummies (baseline hazard).
*** $p<0.001$, ** $p<0.01$, * $p<0.05$, † $p<0.1$.

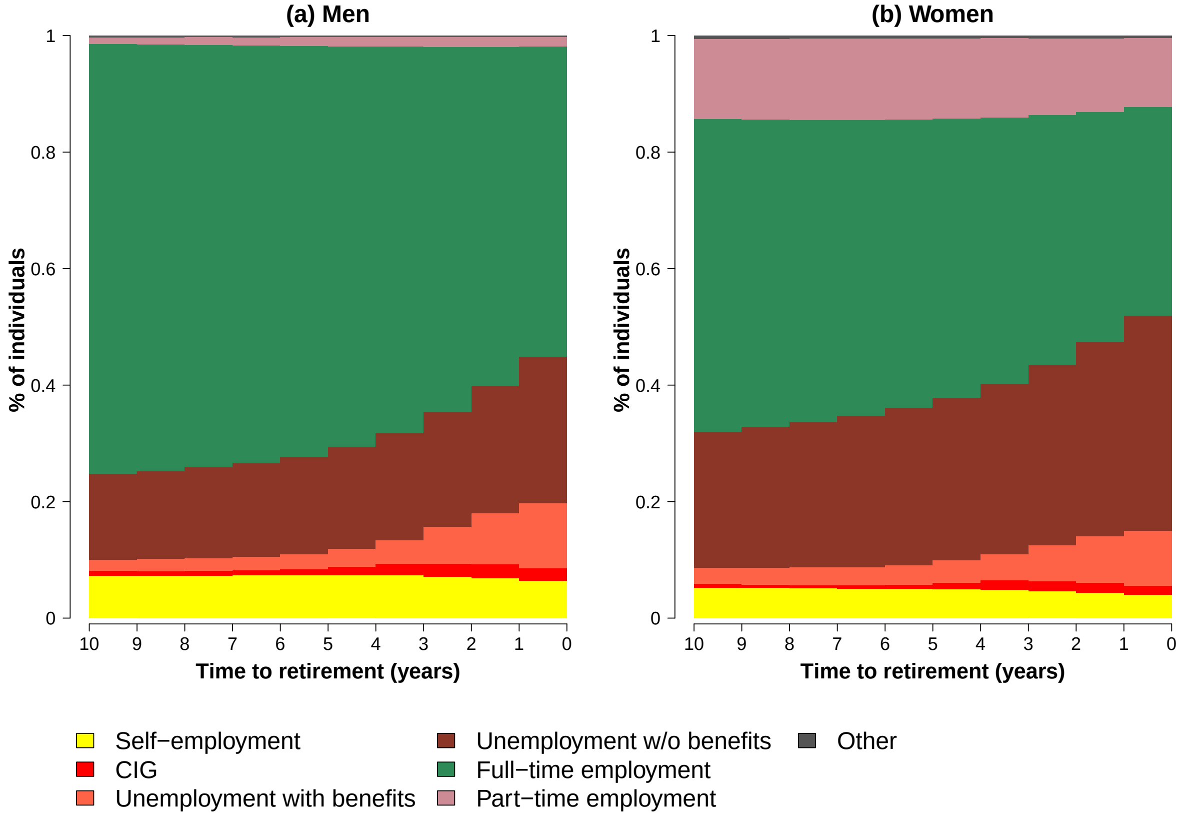
Table 1.9: Robustness checks - health selection

Death	(1) Men	(2) Women
Self-employment	1.075*	1.222**
	(0.035)	(0.087)
Part-time employment	1.155	1.009
	(0.116)	(0.063)
Unempl. w/o benefits	1.152***	1.169***
	(0.030)	(0.053)
Full-time empl./unempl. w/o benefits	1.090*	1.059
	(0.040)	(0.083)
Full-time empl./unempl. with benefits	1.088**	1.156*
	(0.036)	(0.085)
Full-time/Part-time empl.		1.069
		(0.127)
Pension health	1.477***	3.214***
	(0.052)	(0.330)
Self-employment x Pension health	1.218*	0.910
	(0.110)	(0.236)
Parti-time empl. x Pension health	1.124	0.867
	(0.246)	(0.167)
Unempl. w/o benefits x Pension health	1.173**	0.825
	(0.058)	(0.101)
Full-time empl./unempl. w/o benefits x Pension health	1.198**	0.727†
	(0.075)	(0.137)
Full-time empl./unempl. with benefits x Pension health	1.089	1.112
	(0.109)	(0.269)
Full-time/Part-time empl. x Pension health		1.223
		(0.345)
Year of birth	0.976***	0.994
	(0.002)	(0.005)
Age first job	0.986***	1.008†
	(0.003)	(0.005)
Retirement age	1.093***	1.081***
	(0.003)	(0.009)
Work after retirement	0.606***	0.664***
	(0.014)	(0.038)
Survivor pension	1.299***	1.108†
	(0.082)	(0.065)
Occupational status (ref: Blue-collar)		
White-collar	0.846***	1.007
	(0.018)	(0.036)
Manager	0.789***	0.777
	(0.045)	(0.137)
Macro-region of residence (ref: Centre)		
North-East	1.089**	1.048
	(0.029)	(0.053)
North-West	1.018	1.143**
	(0.024)	(0.054)
South	0.875***	0.893*
	(0.021)	(0.050)
Islands	0.853***	0.999
	(0.025)	(0.072)
Observations	1,431,429	775,508
Deaths	16,458	3,921

Notes. Results from complementary log-log models. Dependent variable: death occurrence (0,1). Coefficients are expressed in the exponentiated form (hazard ratios). Robust standard error in parentheses. Pension health is a dummy taking value 1 if the individual is a recipient of disability pension benefits, indemnity pension benefits or social pension benefits. A detailed description of these pension benefits is provided in Appendix A.1. All models include also 18 duration dummies (baseline hazard).
*** p<0.001, ** p<0.01, * p<0.05, † p<0.1.

Figures

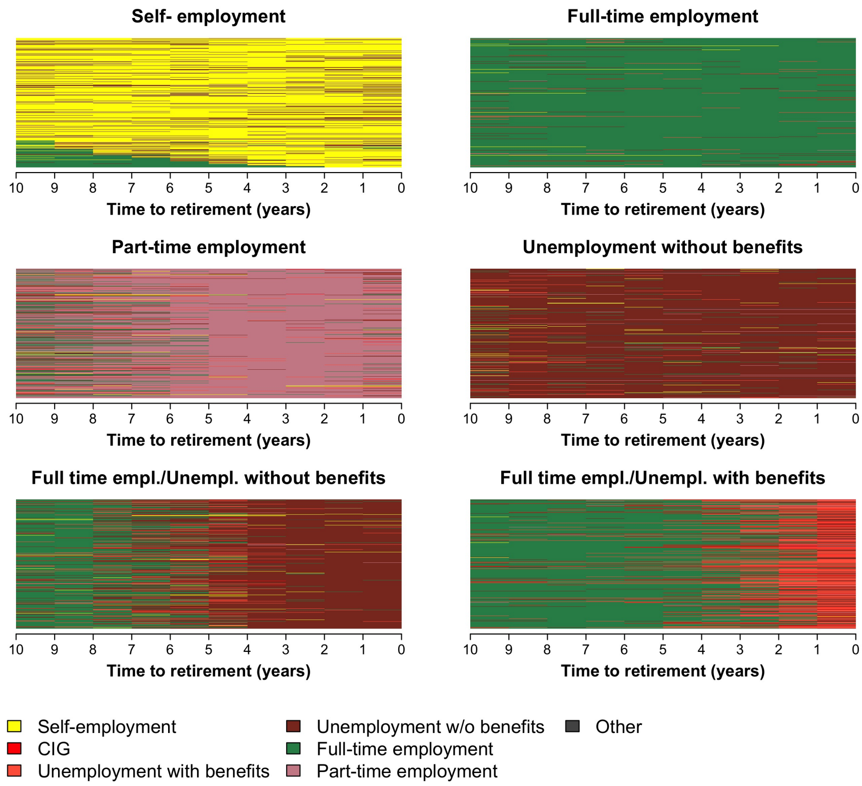
Figure 1.1: Distribution of employment statuses
Ten years prior to retirement



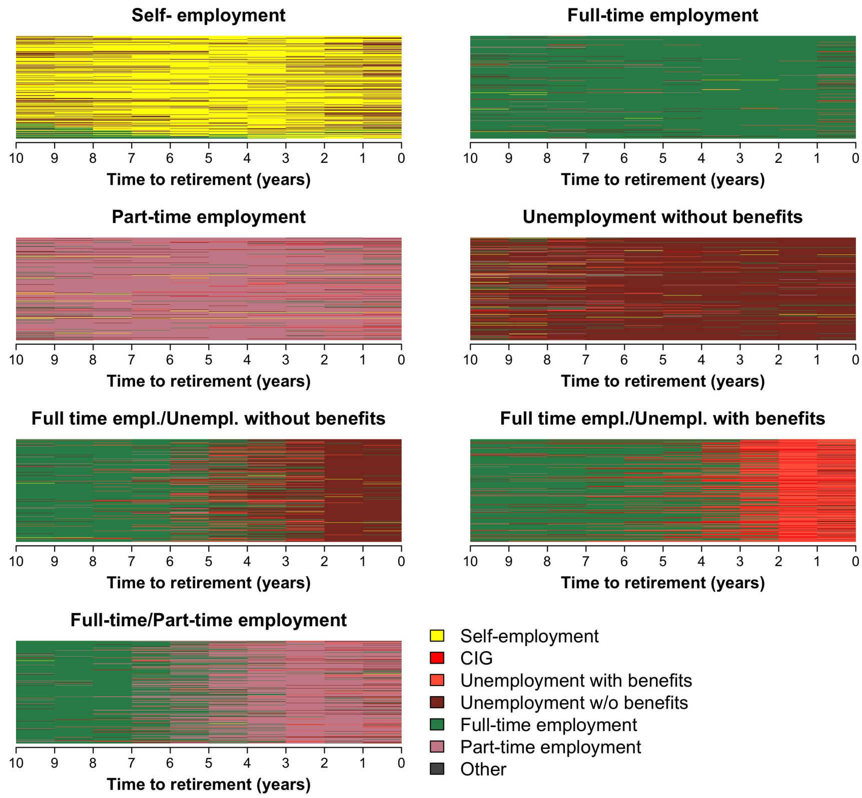
Notes. The horizontal axis measures time to retirement (years). The vertical axis measures the proportion of retirees in each retirement year falling in each employment state. Own elaboration based on INPS LoSai sample.

Figure 1.2: Individual late-career sequences grouped by ideal-type employment trajectories

(a) Men

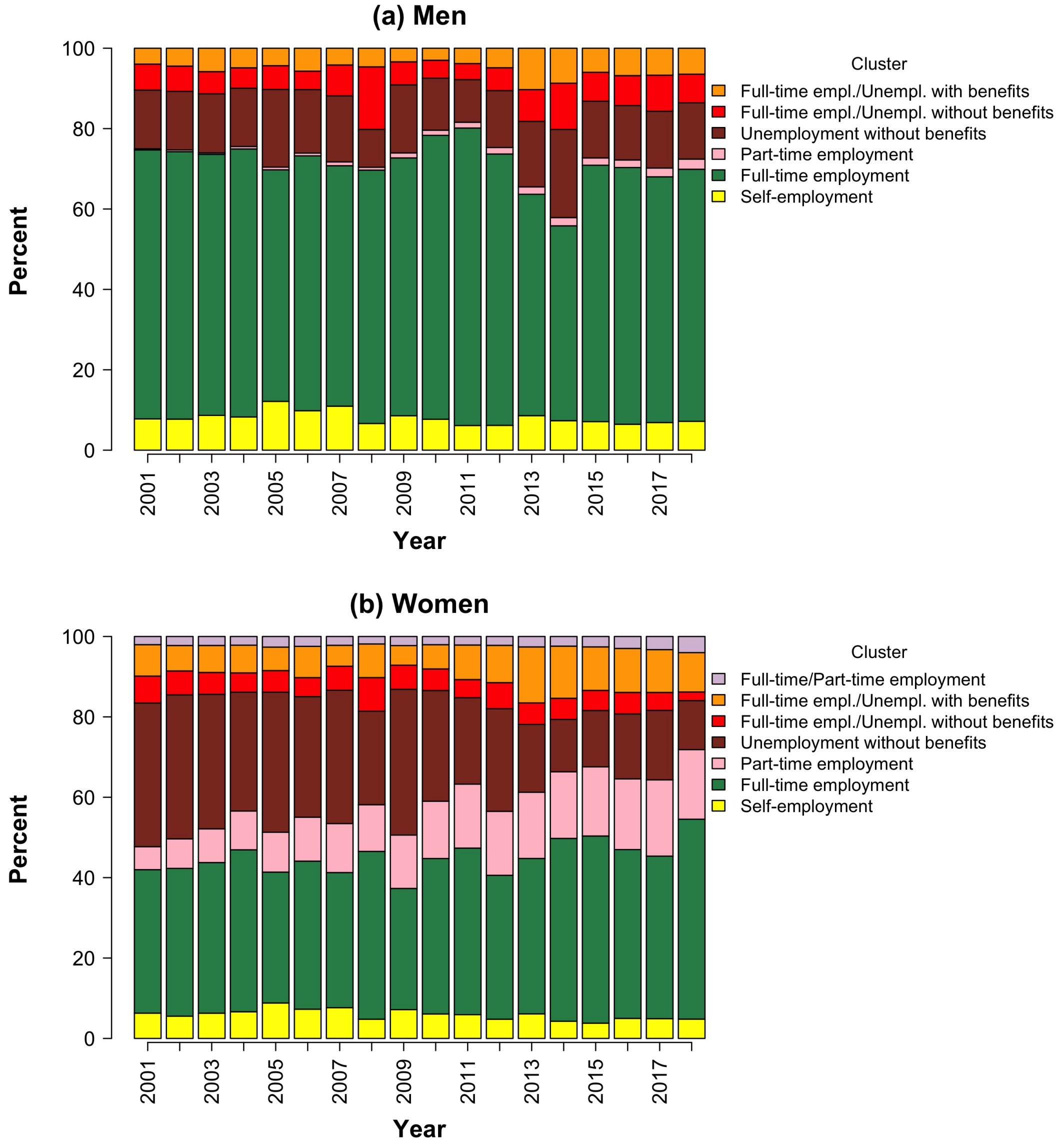


(b) Women



Notes. The horizontal axis measures the time to retirement (years). For each cluster, 1000 representative sequences ordered based on the distance from the most frequent sequence in each cluster are shown. Own elaboration based on INPS LoSai sample.

Figure 1.3: Proportion of ideal-type employment trajectories by retirement year



Notes. The horizontal axis measures the retirement (calendar) year. The vertical axis measures the proportion of retirees in each retirement year falling in the identified clusters.

Chapter 2

Inequalities in mortality by lifetime specific occupation among Italian retirees, 2010-2019

2.1 Introduction

Longevity in low mortality countries is highly stratified by socio-economic characteristics. Regardless of how one measures socio-economic status, whether by education, income or occupational class, life expectancy tends to be shorter as one moves down the socio-economic ladder (Mackenbach et al. 2019; Marmot 2005; Wilkinson & Marmot 2003). While correlated, though, education, income and occupation cannot be used interchangeably. Indeed, if education predicts the ability of turning information into behavioural choices and income proxies the availability of material resources, occupation is more suited to measure social prestige and job control, and to account for exposure to work-specific risks and benefits over one's working life (Cambois et al. 2020; Geyer 2006). Recent research efforts have focused mostly on longevity differentials by income and education, pointing at widening lifespan inequalities along both dimensions in a number of OECD countries (Auerbach et al. 2017; Permanyer et al. 2018; Sasson 2016). Instead, interest in disparities in mortality by occupational class has markedly declined (Katikireddi

[1]This chapter is the result of a joint work with S. Ghislandi. We are thankful to F.C. Billari and to participants to the VisitINPS seminars for useful comments and suggestions. Access to administrative data was kindly granted by the Italian Social Security Institute as part of the VisitINPS Scholar Program. The opinions expressed in this study are those of the authors. They do not purport to reflect the opinions or views of the Italian Social Security Institute or its members.

et al. 2017).

From a policy perspective, though, tracking differences in mortality across occupational groups is highly relevant for setting key programs, ranging from targeted health prevention interventions to equitable retirement policies. For these purposes, detailed and timely assessments of mortality by occupation are needed. With few exceptions (Johnson et al. 1999; Katikireddi et al. 2017), studies from most countries focus on few, broad occupational classes defined over the manual-non manual spectrum. By lumping together rather heterogenous categories, this big-class approach limits the room for assessing health inequalities emerging around specific occupations (Weeden & Grusky 2012). Moreover, available studies on mortality differentials across occupational groups are mostly based on mortality observed among working-age individuals. While crucial for detecting patterns of premature mortality, focusing on working-age populations does not allow to quantify differences in mortality at older ages associated to specific occupational backgrounds. This kind of information is paramount, for instance, to inform policies which aim at balancing the need for raising statutory retirement age with that of ensuring early retirement options for vulnerable categories of workers.

Building on these considerations, in this study we investigate mortality patterns among Italian retirees aged 65-74 by former (lifetime) occupation, defined on the basis of a highly detailed taxonomy encompassing more than thirty occupational categories. For a thorough assessment of the implications of such patterns, we also project life expectancy at 65 for each occupational category. To the best of our knowledge, this study represents the first attempt to explore post-retirement mortality patterns and estimate life expectancy around retirement age by specific lifetime occupation in Italy, and in a low mortality country in general. Previous studies based on broad occupational groups find that lifespan inequalities in Italy are relatively limited when compared to other low-mortality countries (Lallo & Raitano 2018; Leombruni et al. 2015). Italy thus represents an interesting setting for assessing how much insight about the longevity distribution along the occupational dimension can be gained through detailed mortality analyses across specific occupational groups.

2.2 Background

As recurrently documented across a number of countries, individuals belonging to upper non-manual and high-skilled occupational groups tend to live longer than individuals belonging to lower manual and low-skilled occupational groups (Mackenbach et al. 2019; van Raalte et al. 2014). Generally, this kind of evidence comes from studies which classify occupations on a broadly defined basis, with typically less than eight categories (Tanaka et al. 2019). Mortality analyses by specific occupation are rare. A couple of exceptions stand out. In a seminal study based on data from the U.S. National Longitudinal Mortality Study, Johnson et al. (1999) estimate relative all-cause mortality risks among individuals aged 20-64 using a detailed occupational taxonomy, documenting the existence of sizeable heterogeneities in mortality beyond those accounted for by social status, income and education. As suggested by their analysis, high-risk specific occupations include taxi drivers, cooks, and transportation operatives, while low-risk occupations encompass professionals such as lawyers, natural scientists, teachers, engineers, but also farmers. In a more recent work, Katikireddi et al. (2017) analyse patterns of all-cause mortality in the UK among working age-individuals (20-59) across more than sixty occupations based on linked census and mortality records spanning years 1991-2011. They find occupation-specific mortality rates to differ by more than three times between the lowest and highest observed rates in both men and women, excess mortality being concentrated among low-skilled manual occupations such as elementary construction, housekeeping and factory workers.

In Italy, a few studies have investigated the relationship between occupational class, broadly defined, and mortality. Linking 2011 census data with mortality records over

2012-2014, Bertuccio et al. (2018) estimate all-cause and cause-specific mortality rates by occupation-based social class in the Italian working-age population (20-64) using the Erikson–Goldthorpe class schemes whereby occupations are classified into 7 categories.[1] Their analysis documents the existence of substantial heterogeneities among males, mortality for a large number of causes being higher among non-skilled manual workers. Instead, they find limited differences in mortality among working-age women, which are entirely accounted for by adjustments for education. These patterns are consistent with those documented by Leombruni et al. (2015), who find a clear gendered occupational gradient in post-retirement mortality among individuals formerly employed in the private sector, based on four occupational groups (blue-collar workers, white-collar workers, managers, self-employed). Using social security data spanning years 1974-2012, they estimate a gap of about 1.8 years in residual life expectancy at 65 between former blue-collar workers and managers in the case of men, but no tangible differences in the case of women. In a similar vein, Lallo & Raitano (2018) combine social security data with survey data from the Italian 2005 EU-SILC module to estimate life expectancy at 60 by macro-occupational class (employees, self-employed, farmers), adjusted for possible confounders such as education and household economic conditions. Their estimates, based on a mortality follow-up spanning years 2005 through 2009, document a difference of 5 years in remaining life expectancy at 60 between men with opposite socioeconomic statuses. As all these studies rely on relatively broad categorizations of occupational class, they may fail to detect important heterogeneities.

[1] Upper non-manual workers, routine non-manual workers, self-employees, farmers, skilled manual workers, non-skilled manual workers and agricultural labourers.

2.3 Data and Methods

2.3.1 Data sources

We rely on three datasets extracted from the digital archives of the Italian Social Security Institute: the *Comunicazioni Obbligatorie* dataset, the *Casellario Pensioni* dataset, and the *Anagrafica* dataset. The *Comunicazioni Obbligatorie* (COB) dataset, originally provided by the Italian Ministry of Labour, keeps track of all events entailing the creation, cessation, and transformation of job relationships in both the private and public sector in Italy, between 2010 and 2019.[2] For each event, we have information about the beginning and (when relevant) ending date of the job relationship, as well as about the occupational class and the education level of the individual which the event refers to.[3] Occupational class is categorized according to the *Classificazione delle Professioni 2011* (CP2011) taxonomy compiled by the Italian National Institute of Statistics. The CP2011 classification represents the Italian version of the most recent International Standard Classification of Occupations (ISCO-08)[4] and is hierarchically structured, with five-digit occupational codes being the most detailed and one-digit occupational codes the least. The *Casellario Pensioni* dataset reports all pension benefits disbursed by INPS-managed social security schemes between 1995 and 2018. Pension benefits disbursed by INPS fall into four main categories: old-age/seniority pensions (pensions based on previous work contributions), disability pensions (paid to INPS-insured individuals of working age who are temporarily or permanently unable to work due to physical or mental impairment), social disability pensions (paid to all individuals, whose health conditions limit their work capacity completely and on a permanent basis) and social pensions (which include means-tested benefits for poor pensioners and attendance allowances). The *Casellario Pensioni*

[2]It should be noted that the COB dataset does not provide, *per se*, a number of information which are key for the proper reconstruction of individual work histories such as salary, industry/sector of employment, etc.

[3]It is important to note that the accuracy of education level recording in the COB dataset is relatively poor. Indeed, education level tends to vary upon job change for the very same individual. In our analysis, we consider the last education level recorded for the last occupation held prior to retirement.

[4]In CP2011, occupations are classified from very specific classes (5-digit titles) to broad classes (1-digit titles), corresponding to ISCO-08 major groups

dataset provides also information about pensioners' place of residence and marital status. Finally, the *Anagrafica* dataset reports basic demographic characteristics of all individuals appearing in the INPS archives, including gender, month and year of birth, month and year of death (if relevant), month and year of retirement (if relevant). Information in the *Anagrafica* dataset is updated to December 31, 2019. Observations in these three datasets can be matched through unique individual identifiers.

2.3.2 Dataset construction

The original *Comunicazioni Obbligatorie* dataset includes observations relative to 21,240,742 uniquely identified individuals. Given the objective of our study, we focus on job cessations that are plausibly linked to entry into retirement. For this purpose, we keep individuals who experience a job cessation between 2010 and 2018, and who retired in the same period, aged $\leq$70. We restrict our analysis to individuals whose last job relationship prior to retirement lasted at least 5 years. For these individuals, it is highly likely that their last occupation represents a reliable proxy of the occupation they predominantly held throughout their working life. Reassuringly, and consistently with historically life-long employment relationships in the Italian labour market, the average length of the last job relationship for individuals in our final sample is 25 years (Table 2.1). We further drop individuals who do not appear in the *Casellario Pensioni* dataset, as for these individuals we do not have information about place of residence and marital status, nor about reception of disability pension benefits. As further explained below, we analyse mortality differences across occupational groups for individuals aged 65-74. This implies that individuals who die before 65 or who have not turned 65 by December 31, 2018 are further excluded from the dataset. Our final dataset is made up of 620,146 individuals, 361,829 men and 258,317 women. Over the period of analysis (2010-2019), registered deaths amount to 19,092 (14,253 among men and 4,983 among women).

Table 2.2 recapitulates all the steps taken in the construction of the sample, whose main descriptive statistics are reported in Table 2.1. Table 2.3 reports the sex-specific distribution of individuals and deaths over the period of analysis by CP2011 occupational class at the 1-digit and 2-digit level, encompassing eight and thirty-four categories respectively, excluding armed forces. In order to observe a sufficient number of deaths in each group, in our analysis we aggregate two-digit categories, within the same one-digit category, reporting less than 50 deaths over 2010-2019. Because of gender differences in former occupation, groupings differ between men and women (Table 2.4).

2.3.3 Statistical analysis

To explore differences in post-retirement mortality by occupation, we adopt two main methodologies. First, we calculate person-year mortality rates at ages 65-74, stratified by sex and age-adjusted through the indirect method, using sex-specific rates across all occupations as standard (Bessudnov et al. 2011; Johnson et al. 1999). We opt for indirect standardization because the number of deaths in certain age- and sex-specific categories is small (Ahlbom 1993). Methodological details are provided in Appendix B. We decide to restrict the analysis to individuals aged 65-74 because 65 is the age by which most individuals in our period of analysis enter into retirement, and because the number of individuals who turn older than 74 in our dataset by the end of 2019 is negligible. We calculate occupation- and sex-specific mortality rates at the national and at the macro regional level (the latter for occupational classes measured at the 1-digit level only). Second, since we know individuals' time of death, we use Cox proportional hazard regression analysis to study post-retirement mortality differences across occupational groups controlling for possible confounders, for men and women separately. We opt for Cox proportional hazard regression model as it allows to make no assumption about the nature of the hazard function (Cox 1972). Following previous studies (Bessudnov et al. 2011; Lallo & Raitano 2018), we use age as analytic time variable, setting entry time at age 65 or age at retirement, whichever later, and exit time at age 74 or age attained by the end of 2019, whichever earlier. We summarize these choices in a Lexis-type diagram reported in Figure B.1.

We consider two specifications. In the baseline specification, we model the relationship between mortality and occupational class only, stratified by year of birth and year of retirement. The baseline specification looks as follows:

$$h_i(t) = h_0^{\sigma}(t) \times e^{\beta_j Occupation_{ij}} \tag{2.1}$$

where the subscripts i and j indicate individual i and occupation j, respectively, and $h_0^{\sigma}(t)$ is the baseline mortality hazard, stratified by year of birth and year of retirement (σ). In the extended specification, we add controls for factors which may plausibly correlate with occupational class and mortality, including educational level, marital status, macro-region of residence (including residence abroad), reception of disability benefits/social disability benefits. In estimating parameters, we account for both left truncation in age at entry and right censoring. Moreover, as a consequence of using age as time of entry and exit, our dataset presents a large number of ties, i.e. contemporary entry and exit of individuals. We tackle this issue by applying the Efron method, which is particularly suited for handling multiple ties (Efron 1977).

In order to better assess the implications of differential mortality after retirement, it is useful to translate the estimated mortality hazards into metrics providing a reliable measure of residual lifespan. For this purpose, we estimate both partial life expectancy at ages 65-74 and full life expectancy at age 65. Using the Kaplan-Meier method, we first estimate the survivor function for each occupational class between ages 65 and 74, accounting for left-truncation in age at entry and right censoring. Partial life expectancy is then computed as the area below the occupation-specific survival curve, from age

65 to age 74. To estimate full life expectancy, we extrapolate survival curves using two-parameter Brass relational logit model, which is commonly used in the presence of incomplete survival curves (Brass 1971; Wilmoth et al. 2011). The classical Brass relational model posits the existence of a linear relationship between the logits of any two human survival curves. One can therefore obtain complete survival profiles by relating the logits of any incomplete survival curve, Y_x, to the logits of a standard (complete and trustworthy) survival curve, Y_x^s:

$$Y_x = \alpha + \beta Y_x^s \tag{2.2}$$

where α and β are the parameters of the model, estimated via linear regression, and x is the subscript for age. Y_x and Y_x^s are derived directly from the survival curves by applying the following logit transformations:

$$Y_x = \frac{1}{2} ln\Big[\frac{l_x}{1 - l_x}\Big] \tag{2.3}$$

$$Y_x^s = \frac{1}{2} ln\Big[\frac{l_x^s}{1 - l_x^s}\Big] \tag{2.4}$$

where l_x are the values of the incomplete survival curve (in this case, each of the simulated gender-specific survival curves for all occupational groups) and l_x^s are the values of the complete survival curve (in this case, the official survival curve of the Italian population, by gender, certified by the Italian National Institute of Statistics).[5] As further discussed below, a major limitation of this approach is that it assumes that survival profiles prevailing over the observed age range (65-74 in this case) will persist at older ages. Full life expectancies are then calculated as the areas below the complete occupation-specific

[5]We use the survival curves of the Italian population certified by the Italian National Institute of Statistics in 2018 (the most recent at the time of writing).

survival curves, from age 65 to age 119.

2.4 Results

2.4.1 Occupation-specific mortality rates

Table 2.5 and Table 2.6 report the number of observed deaths, person-years and age-standardized mortality rates per 100,000 person-years at ages 65-74 across occupational groups, for men and women respectively. Occupations are classified based on both 1-digit (macro) and 2-digit (micro) occupational codes. In the case of men, we observe decreasing mortality rates across occupations moving from low-skill elementary occupations to high-skill occupations. At the macro-occupational level, lowest mortality is observed among former managers and senior officials (778 deaths per 100,000 person-years [95% CI 721-838]), while highest mortality is observed among male retirees who used to hold an elementary occupation (1548 deaths per 100,000 person-years [95% CI 1482-1615]). Analysis by regions reveals that differences in mortality rates between high-skill and low-skill occupational groups are particularly pronounced in the North of Italy (Table 2.7). Mortality rate in elementary occupations in the North-East and in the South is 73% and 11% higher compared to mortality rate in managers and senior officials, respectively. At the micro-occupational level, lowest mortality rates are reported by former engineers and architects (650 deaths per 100,000 person years [95% CI 494-829]), while highest mortality rates are observed among former labourers in mining, construction and manufacturing (1,629 deaths per 100,000 person years [95% CI 1481-1784]).

In the case of women, differences in mortality rates across occupational groups are considerably less pronounced. At the macro-occupational level, lowest mortality rates are recorded among former professionals (507 deaths per 100,000 person-years [95% CI 479-535]), and highest ones among women formerly employed in elementary occupations (664 deaths per 100,000 person-years [95% CI 615-714]). Contrary to men's case, no significant differences emerge in occupation-mortality rates at the regional level among female retirees, while substantial heterogeneity emerges within macro-occupational groups. For instance, focusing on professional occupations, women formerly employed as teaching and research professionals display significantly lower mortality (475 deaths per 1000,000 person-years [95% CI 443-508]) compared to women formerly employed as legal, social and cultural professionals (616 deaths per person-years [95% CI 553-682]). It is also worth stressing that women with specific high-skill occupational backgrounds (e.g. managers) display relatively high mortality rates, while women with specific low-skill occupational backgrounds (e.g. cleaners and helpers) display relatively low mortality rates.

2.4.2 Cox proportional hazard regression models

Table 2.8 and Table 2.9 report results from Cox proportional hazard regression models for men, where occupations are classified based on 1-digit (macro) and 2-digit (micro) occupational codes respectively. Coefficients are expressed in the exponentiated form (hazard ratios). When stratifying for year of birth and year of retirement only, we document a clear occupational gradient in mortality at ages 65-74 across macro-occupational groups (Table 2.8, Column 1). Compared to the reference group (Clerical support workers), individuals in upper non-manual occupations face substantially lower mortality risk between 65 and 74. Managers and senior officials display the lowest mortality risk (-35%), followed by professionals (-29%) and technicians (-13%). On the contrary, individuals in lower manual or unskilled groups face substantially higher mortality risk. Male retirees holding a background in elementary occupations face the highest mortality risk (+17%), followed by plant and machine operators/assemblers (+13%) and craft and related trade workers and skilled workers in agriculture, forestry and fishery (+11%). When adding controls for possible confounders (Table 2.8, Column 1), estimated hazards change slightly in magnitude, but occupational background remains a powerful determinant of males' post-retirement mortality. Cox proportional hazard regression analysis based on micro-occupational groups yields broadly consistent results, allowing to identify specific high- and low-risk occupational profiles (Table 2.9). Looking at the extended specification (Table 2.9, Column 3), male retirees holding an occupational background in engineering, architecture and similar professions display the lowest mortality risk (-28%) compared to the reference group (General and keyboard clerks), followed by former managing directors and chief executives (-22%). Comparatively high-risk profiles include unskilled sales workers, cleaners and helpers (+77%), labourers in mining, construction and manufacturing (+22%), and assemblers (+17%). While occupational inequalities in mortality are the core object of this study, there is also some interest in the estimated associations between mortality and control variables. Focusing on Table 2.9, higher education comes with lower post-retirement mortality risk, all other things equal. Men with tertiary education (university degree) face a 17% lower mortality hazard compared to men with primary education. No statistically significant differences emerge, instead, for those holding secondary education. Marital status is a remarkably strong and consistent predictor of post-retirement survival. *Ceteris paribus*, widowed, separated/divorced, and unmarried men have all higher post-retirement mortality risk compared to married men: +25%, + 38% and +59%, respectively. Macro-region of residence is also significantly associated to mortality: for men living abroad and in the North-East and North-West of Italy, the mortality hazard ratios are respectively 30%, 16% and 8% higher compared to men living in the Centre, while no statistically significant differences emerge for those residing in the South-Islands and abroad, all other things equal. Finally, as one may expect, men who receive disability pension benefits face substantially higher mortality risk compared to non-recipients, *ceteris paribus* (41% higher in the case of ordinary disability benefits and 847% higher in the case of social disability benefits).

Results for women are reported in Table 2.10 and Table 2.11. In this case, we find limited evidence of occupation gradient in mortality over the considered ages across macro-and micro-occupational groups, in both the baseline and extended specifications. In fact, hazard ratios, albeit imprecisely estimated, suggest that some categories of women at the very top of the occupational hierarchy, such as managing directors and chief executives, may actually face higher post-retirement mortality risk compared to the category of reference (Table 2.11). On the contrary, women belonging to manual or unskilled occupations, such as cleaners and helpers, display lower post-retirement mortality compared to the reference group. It is worth noting that the lack of a clear occupational gradient in mortality among female retirees aged 65-74 is consistent with the lack of a clear gradient over the educational dimension. Indeed, women with secondary and tertiary education do not face significantly lower mortality risk compared to women with primary education, other things equal. Instead, marital status is a strong predictor of mortality in the case of female retirees too. Focusing on Table 2.11, widowed, separated/divorced and never married women face mortality hazard ratios which are 26%, 51% and 78% higher compared to married women, ceteris paribus. The same holds for disability benefits: women who receive ordinary and social disability benefits are exposed to a post-retirement mortality risk which is 70% and >1300% higher than non-recipients, other things equal. Finally, macro-regional disparities in post-retirement survival among women are qualitatively and quantitatively analogous to those recorded in the case of men.

The reliability of the results presented in this section depends on the validity of the main assumption of the Cox model, that is the proportionality of hazards. We check this assumption by examining Schoenfeld residuals after fitting the baseline and the extended models, for men and women separately. Results for our main explanatory variables, i.e. the occupational category dummies, are largely reassuring. In all models, we find the proportional hazard assumption to hold for all occupational category dummies, for both men and women (Table B.1 to Table B.4 in Appendix B).

2.4.3 Life expectancy estimates

We now examine how mortality differentials at ages 65-74 by former occupation translate into lifespan differentials. Figure 2.1 and Figure 2.2 display partial life expectancies between ages 65-74 for men, together with the respective 95% confidence intervals. Consistently with results delivered by Cox proportional hazards regression models, we document a fairly clear occupational gradient in residual lifespans. Partial life expectancies across macro-occupational groups range from 8.4 years for male retirees holding a background in elementary occupations and as machine operators/assemblers, to 8.7 for former managers and senior officials (Figure 2.1). Looking at specific occupations, the highest partial life expectancy is recorded by former engineers, architects and similar professionals, followed by legislators and senior officials, and by managing directors and chief executives, while the lowest is displayed by numerical and material recording clerks, assemblers and labourers in mining, construction, and manufacturing. When extrapolating survival curves to obtain full life expectancies at 65, the gap between the bottom and the top of the lifespan distribution clearly widens. Between former managers and senior officials (e_{65}=20.27) and plant machine operators/assemblers (e_{65}=16.82) there is a gap in life expectancy of about 3.4 years (Figure 2.3). Disparities by specific occupational class are even more pronounced. Indeed, at age 65 former engineers, architects and similar professionals (e_{65}=20.76) can expect to live 5 years longer than former protective service workers (e_{65}=15.73) (Figure 2.4).

Results for women are displayed in Figure 2.5 to Figure 2.8. In line with results yielded by Cox proportional hazard regression models, there is limited occupational gradient in female retirees' partial and full life expectancies. The distribution of residual lifespans across occupational groups is markedly narrow, with hardly statistically discernible differences across most groups, as witnessed by overlapping confidence intervals. Focusing on macro-occupational groups, partial life expectancies at 65-74 range from 8.66 in the case of plant and machine operators/assemblers to 8.78 in the case of professionals (Figure 2.5). In the case of specific occupations, teaching and research professionals boast the highest partial life expectancy (8.79) and clerical support workers the lowest (8.63) (Figure 2.6). Looking at full life expectancies across macro groups, at 65 females with an occupational background as managers and senior officials can expect to live 2.4 years longer than former plant and machine operators/assemblers (Figure 2.7). At the micro level, the lack of a clear occupational gradient becomes particularly visible. Indeed, female retirees holding a background in lower manual and unskilled occupations, such as cleaners and helpers (e_{65}=21.99) and unskilled sales workers (e_{65}=21.21), can expect to live longer than former managers (e_{65}=20.81) or legal, cultural and social professionals (e_{65}=20.46). It is worth noticing that in the case of women, tangible disparities in mortality emerge within macro-occupational groups themselves. For instance, legislators and senior officials (e_{65}=22.17) boast an advantage in life expectancy at 65 of about 1.4 years compared to managers (e_{65}=20.81), a group which includes managing directors, chief executives, and professional services managers.

All estimates of partial and full life expectancies by sex and occupational group, along with their respective confidence intervals, are reported in Table B.5 to Table B.12 in Appendix B.

2.5 Discussion

Mortality patterns by lifetime occupation among Italian retirees differ substantially between men and women. In the case of men, we find that post-retirement mortality follows a neat occupational gradient, which holds also when accounting for potential confounders, including education, marital status, macro-region of residence and physical/mental impairment proxied by the reception of disability benefits. Specifically, we observe increasing mortality risk moving from highly qualified, non-manual occupations (such as engineers and architects, legislators and public senior officials, managing directors and chief executives) to manual, low-skilled and generally labour-intensive occupations (such as labourers in mining, construction and manufacturing, unskilled sales workers and assemblers). Our projections for life expectancy at 65 suggest that in Italy male retirees with a background in specific low-risk occupational categories enjoy an advantage of about 4-5 years compared to those with a background in specific high-risk categories. Such disparities in mortality and longevity are substantially larger than those documented by previous studies employing broader categorizations of occupational profiles (Leombruni et al. 2015).

While occupation-based mortality inequalities among male retirees are observed throughout Italy, our estimates of mortality rates at the macro-regional level suggest that such inequalities are particularly pronounced in the North of Italy. These patterns are consistent with findings by Ardito et al. (2021) who document sizeable regional variation in life expectancy inequalities at 65 among men in Italy by socio-economic status, the gap being significantly larger in Northern regions as compared to Southern regions. Possible explanations include greater income inequalities and higher cost of living in the North as compared to the South. However, it is worth stressing that our data do not allow to capture mortality patterns of individuals (formerly) involved in the informal labour market. Since the incidence of the shadow economy is larger in Southern regions (Italian National Institute of Statistics 2020), and low-skill workers have greater chances to be in informal employment (Viviani 2010), our analysis may underestimate mortality inequalities along the occupational gradient among male retirees in the South of Italy.

Contrary to men's case, we document limited occupation-based mortality gradient among female retirees. This result is consistent with findings by Bertuccio et al. (2018), who observe no significant mortality differences among Italian women aged 20-64 by occupation-based social class when taking educational level into account. In fact, our analysis suggests that women with a background in specific upper non-manual and high-skilled occupations, such as former top managers, may face analogous, if not worse, survival chances into old-age compared to females previously employed in lower manual or elementary occupations. Such patterns of reversed mortality gradient among female retirees are in line with findings by Costa et al. (2017) based on the population of Turin, in northwest Italy. Using census data linked to mortality records, they show that the mortality profile of women in managerial and entrepreneurial careers is similar to that of women in skilled blue-collar occupations. A possible explanation relates to behavioural factors such as smoking which stood as hallmarks of women's emancipation for the considered cohorts (Di Novi & Marenzi 2019).

This study presents some limitations. First, it analyses mortality dynamics for individuals who survive to 65, at least. As a consequence, the working sample may suffer from an under-representation of occupational categories which are systematically exposed to higher risk of premature mortality. As the latter tends to concentrate disproportionately in lower occupational strata (Lewer et al. 2020), our study may fail to properly characterize occupation-driven mortality differentials in the Italian senior population. An alternative approach to avoid such issue, which could be implemented in future works, is to employ time from retirement, rather than age, as analytic time variable, and estimate occupation-specific retirement expectancies (i.e. average years spent in retirement), rather than life expectancies. Second, our study is based on social security data which are representative of the retired population in Italy made up of former private and public employees. As such, it does not examine mortality patterns of individuals with a background in self-employment or without a formal employment background *tout court*, such as workers involved in the informal labour market or homemakers. Third, it defines occupational class based on the last job held prior to retirement. The latter could be a poor proxy of occupational background for individuals with highly discontinuous occupational trajectories. We sought to attenuate this concern by focusing on individuals whose last job prior to retirement lasted at least 5 years, i.e. individuals whose last job can plausibly represent a good measure of lifelong occupational class. From this perspective, the average duration of last job relationship in our sample (25 years) is strongly reassuring. Fourth, it makes heavy use of extrapolation in the estimation of full life expectancies. Estimated mortality patterns across occupational groups beyond age 74 are mostly driven by patterns observed between ages 65-74. This approach may lead to over-estimation of differences in life expectancy across occupational groups if the occupation-mortality gradient attenuates, or even reverses, at older ages, due to frailty-related dynamics (Vaupel & Yashin 1985). Fifth, it focuses on all-cause mortality only. As such, it remains silent about potential mechanisms linking post-retirement survival chances to occupational background which could be inferred by cause-specific mortality patterns.

This study has also two major strengths. First, it is based on administrative data which allow to retrieve reliable information on occupation held prior to retirement. As such, our definition of occupational background does not suffer from self-reporting issues affecting survey and census data. In addition, the richness of our data allows us to explore mortality patterns by former occupational category in greater detail than usually possible. To the best of our knowledge, ours is the first attempt to translate mortality profiles into remaining life expectancies around retirement by specific occupational background.

This study is particularly relevant for policymakers in Italy, and in countries confronting challenging reform needs to meet rising pressures on social security systems posed by increased longevity. While raising statutory retirement age is generally presented as an unavoidable choice, there are mounting concerns that this kind of measures, if applied homogenously, may penalize categories of workers facing unfavourable survival profiles compared to population average, advantaging those with better survival chances (Ayuso et al. 2017; Lozano & Solé-Auró 2021). Moreover, since longevity tends to be stratified by occupation, these measures may amplify intra-generational inequalities emerging over individuals' working life. In the light of these considerations, our study confirms the importance of policy measures aimed at easing access to pension benefits for individuals belonging to high-risk occupational groups. In the case of Italy, it points at the need for extending the official taxonomy of 'demanding' jobs, giving access to early retirement options and to subsidized benefits for early labor market exit,[6] to specific categories which are currently excluded, such as protective service workers. It also points at the need to account for occupation-driven mortality inequalities in the definition of official coefficients used for pension benefits calculation under notional defined contribution (NDC) pension rules.[7] Future works may seek to extend the present analysis by investigating inequalities in healthy or disability-free life expectancy at retirement by specific occupational background. This kind of studies may be extremely valuable in setting differentiated exit paths from the labour market accounting for both residual lifespan and work ability.

[6] The most relevant subsidized benefit scheme for early labor market exit is called *Anticipo pensionistico Sociale*, known as APE sociale. It was introduced in 2017 and is targeted, among the others, to senior workers with an occupational background in 'demanding' occupations, aged 63 years old and who have accumulated between 30 and 36 years of contributions.

[7] Under the notional defined contribution (NDC) pension regime, pension contributions are (fictitiously) accumulated in an individual fund, and are re-evaluated in line with a moving average of GDP growth. Pension benefits are then computed by multiplying the re-evaluated contributions by a coefficient which depends on remaining life expectancy at retirement. Such coefficients are neutral with respect to gender and other relevant socio-economic characteristics, but they are periodically updated to account for changes in official life expectancy projections. NDC pension rules were introduced in 1995 by the Dini reform, but the phase-in period was set to be very long. Indeed, workers with at least 18 years of contributions as of December 1995 were fully unaffected by the reform. Instead, those with a shorter contributory record were to be affected on a *pro rata* basis, the weight of DB depending on the ratio between pre-1995 to the overall contribution period upon retirement. In 2011, the Fornero reform accelerated the transition to full NDC rules, introducing a *pro-rata* contribution for all workers starting from January 1, 2012. In other words, all pensions awarded from this date onward have an NDC component, regardless of the 18-year contribution period mentioned above.

Tables Chapter 2

Table 2.1: Descriptive statistics

Variable	Mean	SD	Min	Max
Age retirement	62.7	2.8	55.0	70.0
Length last job relationship	25.0	11.7	5.0	50.00
Binary variables (%)				
Primary education	17.0	37.0		
Secondary education	65.0	48.0		
Tertiary education	19.0	39.0		
Married	78.0	50.0		
Widow	8.0	28.0		
Separated/Divorced	6.0	25.0		
Never married	8.0	27.0		
Centre	23.0	42.0		
North-East	19.0	39.0		
North-West	24.0	43.0		
South	23.0	42.0		
Islands	12.0	32.0		
Abroad	1.0	8.0		
Disability pension	1.0	9.0		
Social disability pension	2.0	15.0		

Table 2.2: Steps taken in the dataset contruction

	N individuals	Dropped
Initial dataset	21,240,742	
Keeping N with demografic information	21,221,344	19,398
Keeping N retired between 2010-2018	1,523,039	19,698,305
Keeping N with occupational class information (except armed forces)	1,480,953	42,086
Keeping N with last contract lasting $\leq$ 50 years	1,480,176	777
Keeping N with last contract lasting > 5 years	1,187,904	292,272
Keeping N retiring $\leq$ 70	1,173,217	14,687
Keeping N surviving to 65	1,158,895	14,322
Keeping N who reach 65 by the end of 2018	624,281	534,614
Keeping N appearing in Pensioni Casellario	620,146	4,135
Final dataset	620,146	
Men	361,829	
Women	258,317	

Table 2.3: Individuals and deaths by occupational class and sex

Occupation (CP2011, 1digit)	Occupation (CP2011, 2digit)	N individuals		N deaths	
		F	M	F	M
Managers & senior officials	Legislators & senior officials	3647	9203	70	323
	Managing directors & chief executives	1646	10601	39	299
	Professional services managers	426	1786	10	53
Professionals	Engineers, architects & similar professions	141	2647	2	58
	Health professionals	4071	13650	46	418
	Legal social & cultural professionals	18307	25282	354	902
	Life science professionals	1598	1904	26	66
	Science professionals	446	3544	5	88
	Teaching & research professionals	51818	18171	826	610
Technicians	Business & administration technicians	15585	28813	334	1016
	Life science technicians	16535	11231	301	413
	Public service technicians	13037	4042	327	195
	Science & engineering technicians	1615	23282	28	727
Clerical support workers	Customer service clerks	6139	7355	109	287
	General & keyboard clerks	42837	41786	760	1637
	Numerical & material recording clerks	1683	4512	40	167
	Other clerifcal support workers	2485	5982	30	179
Service & sales workers	Personal care workers	3847	1485	58	67
	Personal service workers	6530	3438	129	196
	Protective service workers	6580	6343	138	268
	Sales workers	8145	7665	139	320
Craft & related trade workers, skilled agricultural, forestry & fishery workers	Electrical & electronic trades workers	763	19230	17	751
	Food processing wood working garment & related trade workers	6119	6313	83	294
	Handicraft & printing workers	1044	2646	16	88
	Mining building & related trade workers	5708	16864	132	867
	Skilled agricultural forestry & fishery workers	274	2831	6	173
Plant & machine operators, assemblers	Assemblers	6332	12511	99	476
	Drivers & mobile plant operators	252	21607	2	929
	Machine operators in agricultural & forestry	373	924	7	36
	Stationary plant operators	945	7987	16	273
Elementary occupations	Cleaners & helpers	6446	706	120	36
	Labourers in mining construction manufacturing	2067	8744	49	444
	Unskilled agricultural forestry & fishery	3400	4504	117	278
	Unskilled sales workers	17476	24240	404	1319

Table 2.4: Grouping of occupational 2-digit CP2011 occupational categories

WOMEN	
Grouping	2-digit CP2011 occupational categories
Managers	Managing directors and chief executives, Professional services managers
Other professionals	Engineers architects and similar professions, Health professionals, Life science professionals, Science professionals (mathematics, computer science, chemistry, physics, biology)
Other clerical support workers	Numerical and material recording clerks, Other clerical support workers
Other craft and related trade workers	Electrical and electronic trades workers, Food processing wood working garment and related trade workers, Electrical and electronic trades workers, Food processing wood working garment and related trade workers, Handicraft and printing workers, Skilled agricultural forestry and fishery workers
Plant and machine operators	Drivers and mobile plant operators, Machine operators in agricultural and forestry, Stationary plant operators
Unskilled workers in agriculture, forestry, fishery, mining, construction, manufacturing	Labourers in mining construction manufacturing, Unskilled agricultural forestry and fishery
MEN	
Grouping	2-digit CP2011 occupational categories
Machine operators, drivers & mobile plant operators	Drivers and mobile plant operators, Machine operators in agricultural and forestry

Table 2.5: Mortality rate at 65-74 by former occupation - Men

Occupation	N. of deaths	Person-years	Mortality rate†
Managers and senior officials	675	83669	778 [721-838]
Legislators and senior officials	323	39401	766 [685-852]
Managing directors and chief executives	299	38232	773 [688-863]
Professional services managers	53	6036	913 [684-1176]
Professionals	2142	246795	854 [818-890]
Engineers, architects and similar professions	58	9134	650 [494-829]
Teaching and research professionals	610	74872	789 [728-853]
Health professionals	418	51196	791 [717-869]
Science professionals (maths, computer science, chemistry, physics, biology)	88	10758	851 [682-1038]
Life science professionals	66	7386	872 [674-1095]
Legal, social and cultural professionals	902	93449	961 [899-1025]
Technicians	2351	223471	1063 [1020-1106]
Business and administration technicians	1016	97441	1048 [985-1114]
Science and engineering technicians	727	70742	1064 [988-1143]
Life science technicians	413	38375	1084 [982-1191]
Public service technicians	195	16913	1094 [945-1252]
Clerical support workers	2270	184757	1265 [1214-1318]
General and keyboard clerks	1637	133567	1253 [1193-1315]
Customer service clerks	287	23858	1236 [1097-1383]
Numerical and material recording clerks	167	11565	1551 [1325-1795]
Other clerical support workers	179	15767	1209 [1038-1393]
Service and sales workers	851	61843	1371 [1280-1465]
Personal care workers	67	5156	1245 [964-1560]
Protective service workers	268	20856	1300 [1149-1461]
Sales workers	320	23918	1340 [1197-1490]
Personal service workers	196	11918	1607 [1390-1840]
Craft & related trade workers, skilled agricultural, forestry & fishery workers	2173	148263	1458 [1397-1520]
Handicraft and printing workers	88	6981	1308 [1049-1595]
Mining, building and related trade workers	867	59965	1391 [1300-1485]
Skilled agricultural, forestry and fishery workers	173	11021	1490 [1277-1721]
Electrical and electronic trades workers	751	51198	1526 [1419-1637]
Food processing, wood working, garment &other craft & related trades workers	294	19098	1533 [1363-1713]
Plant and machine operators, assemblers	1714	120319	1484 [1415-1555]
Stationary plant operators	273	21589	1336 [1182-1499]
Machine operators in agricultural/food industry, drivers & mobile plant operators	965	67111	1478 [1387-1573]
Assemblers	476	31619	1599 [1458-1745]
Elementary occupations	2077	129789	1548 [1482-1615]
Unskilled agricultural, forestry and fishery workers	278	18819	1343 [1190-1506]
Unskilled sales workers, cleaners and helpers	1355	84009	1572 [1489-1656]
Labourers in mining, construction, manufacturing	444	26961	1629 [1481-1784]

† Mortality rates were standardized by age using the indirect method with sex-specific rates across all occupations used as standard. Rates are per 100,000 person-years

Table 2.6: Mortality rate at 65-74 by former occupation - Women

Occupation	N. of deaths	Person-years	Mortality rate†
Managers and senior officials	119	19509	596 [494-708]
Legislators and senior officials	70	12802	531 [414-662]
Managers	49	6707	724 [536-941]
Professionals	1259	248521	507 [479-535]
Science, engineers, architects, life science, health professionals	79	17520	463 [366-570]
Teaching and research professionals	826	172964	475 [443-508]
Legal, social and cultural professionals	354	58037	616 [553-682]
Technicians	990	163523	593 [556-630]
Public service technicians	327	58283	524 [469-582]
Business and administration technicians	334	52157	631 [565-701]
Science, engineering, life science technicians	329	53083	634 [567-704]
Clerical support workers	939	148902	650 [609-692]
General and keyboard clerks	760	121530	640 [596-687]
Customer service clerks	109	17326	665 [546-795]
Other clerical support workers	70	10046	739 [576-923]
Service and sales workers	464	76068	617 [562-674]
Personal care workers	58	10484	581 [441-741]
Protective service workers	138	21884	605 [508-710]
Personal service workers	129	21312	615 [514-726]
Sales workers	139	22388	647 [544-759]
Craft and related trade workers, skilled agricultural, forestry and fishery workers	254	42156	623 [549-702]
Mining, building and related trade workers	132	20033	662 [553-779]
Other craft and related trade workers	122	22123	587 [487-695]
Plant and machine operators, assemblers	124	19474	692 [575-819]
Plant and machine operators	25	4059	672 [434-961]
Assemblers	99	15415	697 [566-841]
Elementary occupations	690	100112	664 [615-714]
Cleaners and helpers	120	20503	568 [471-674]
Unskilled workers in agriculture, forestry, fishery, mining, construction, manufacturing	166	24534	617 [527-715]
Unskilled sales workers	404	55075	720 [651-792]

† Mortality rates were standardized by age using the indirect method with sex-specific rates across all occupations used as standard. Rates are per 100,000 person-years

Table 2.7: Mortality rate at 65-74 by former occupation and macro-region

	Men				
Occupation (macro)	North-East	North-West	Centre	South	Islands
Managers and senior officials	638 [512-778]	779 [667-900]	888 [773-1012]	794 [669-931]	655 [492-841]
Professionals	839 [751-932]	848 [765-935]	842 [770-916]	875 [808-944]	831 [734-933]
Technicians	1036 [930-1148]	1107 [1011-1208]	1014 [931-1101]	1077 [996-1161]	1097 [978-1222]
Clerical support workers	1003 [875-1139]	1317 [1194-1445]	1232 [1128-1341]	1405 [1310-1504]	1193 [1062-1331]
Service and sales workers	1520 [1264-1799]	1551 [1332-1785]	1310 [1137-1496]	1340 [1172-1518]	1178 [965-1412]
Craft and related trade workers, skilled agricultural, forestry and fishery workers	1496 [1348-1651]	1766 [1628-1909]	1426 [1291-1567]	1350 [1240-1465]	1141 [993-1299]
Plant and machine operators, assemblers	1705 [1520-1902]	1683 [1531-1843]	1381 [1233-1538]	1398 [1270-1531]	1229 [1058-1412]
Elementary occupations	1742 [1524-1974]	1640 [1468-1821]	1601 [1436-1774]	1532 [1426-1641]	1385 [1250-1528]
	Women				
	North-East	North-West	Centre	South	Islands
Managers and senior officials	517 [282-824]	838 [580-1143]	591 [420-792]	437 [254-670]	591 [330-929]
Professionals	503 [429-583]	545 [467-628]	524 [468-583]	501 [453-552]	461 [399-527]
Technicians	599 [508-698]	636 [548-730]	610 [539-685]	589 [519-664]	507 [423-600]
Clerical support workers	714 [612-823]	697 [613-786]	639 [561-722]	642 [551-739]	493 [396-600]
Service and sales workers	569 [465-685]	706 [601-820]	581 [478-694]	690 [521-882]	534 [355-750]
Craft and related trade workers, skilled agricultural, forestry and fishery workers	506 [385-643]	723 [570-893]	583 [455-727]	810 [570-1093]	591 [282-1015]
Plant and machine operators, assemblers	897 [659-1171]	670 [499-867]	373 [192-615]	813 [404-1365]	591 [56-1695]
Elementary occupations	656 [545-777]	768 [664-879]	600 [506-701]	646 [540-761]	615 [487-758]

† Mortality rates were standardized by age using the indirect method with sex-specific rates across all occupations used as standard. Rates are per 100,000 person-years. 95% confidence intervals in parentheses.

Table 2.8: Cox PH regressions - Men (macro-occupational groups)

	(1)	(2)
Occupation (ref: Clerical support workers)		
Managers and senior officials	0.643***	0.847***
	(0.028)	(0.040)
Professionals	0.713***	0.908**
	(0.022)	(0.031)
Technicians	0.867***	0.978
	(0.026)	(0.029)
Service and sales workers	1.078†	1.092*
	(0.044)	(0.044)
Craft and related trade workers, skilled agricultural, forestry and fishery workers	1.106***	1.153***
	(0.034)	(0.036)
Plant and machine operators, assemblers	1.132***	1.208***
	(0.037)	(0.040)
Elementary occupations	1.174***	1.147***
	(0.036)	(0.036)
Marital status (ref: married)		
Widow		1.251***
		(0.046)
Separated/Divorced		1.386***
		(0.046)
Never married		1.577***
		(0.043)
Region of residence (ref: Centre)		
Abroad		1.290*
		(0.139)
Islands		0.968
		(0.029)
North-East		1.081**
		(0.031)
North-West		1.160***
		(0.030)
South		1.019
		(0.025)
Education (ref: Primary)		
Secondary education		0.985
		(0.022)
Tertiary education		0.837***
		(0.031)
Social disability pension		8.451***
		(0.177)
Disability pension		1.400***
		(0.085)
Observations	361,829	361,829

Results from Cox proportional hazards models. Results are presented in the exponentiated form (hazard ratios). Standard errors in parentheses. *** p<0.001, ** p<0.01, * p<0.05, † p<0.1.

Table 2.9: Cox PH regressions - Men (micro-occupational groups)

	(1) Model 1	(2) Model 2
Occupation (ref: General and keyboard clerks)		
Legislators and senior officials	0.629***	0.853*
	(0.039)	(0.057)
Managing directors and chief executives	0.630***	0.785***
	(0.040)	(0.051)
Professional services managers	0.767†	0.834
	(0.107)	(0.118)
Science professionals (mathematics, computer science, chemistry, physics, biology)	0.680***	0.822†
	(0.075)	(0.091)
Engineers, architects and similar professions	0.563***	0.718*
	(0.075)	(0.098)
Life science professionals	0.741*	1.079
	(0.093)	(0.140)
Health professionals	0.664***	0.921
	(0.037)	(0.059)
Legal, social and cultural professionals	0.795***	0.900*
	(0.033)	(0.038)
Teaching and research professionals	0.650***	0.832***
	(0.031)	(0.044)
Science and engineering technicians	0.866**	0.945
	(0.039)	(0.043)
Life science technicians	0.877*	0.983
	(0.049)	(0.055)
Business and administration technicians	0.853***	0.932†
	(0.034)	(0.038)
Public service technicians	0.860*	0.928
	(0.065)	(0.071)
Customer service clerks	0.977	0.804***
	(0.063)	(0.052)
Other clerical support workers	0.917	0.936
	(0.073)	(0.075)
Numerical and material recording clerks	1.075	1.024
	(0.088)	(0.084)
Sales workers	1.048	1.046
	(0.064)	(0.064)
Personal service workers	1.237**	1.190*
	(0.094)	(0.091)
Personal care workers	1.008	0.980
	(0.126)	(0.123)
Protective service workers	1.017	1.004
	(0.067)	(0.066)
Mining, building and related trade workers	1.044	1.077†
	(0.044)	(0.046)
Electrical and electronic trades workers	1.159***	1.173***
	(0.052)	(0.053)
Handicraft and printing workers	1.000	0.999
	(0.110)	(0.110)
Skilled agricultural, forestry and fishery workers	1.167†	1.108
	(0.094)	(0.089)
Food processing, wood working, garment and other craft and related trades workers	1.133*	1.130†

Continued on next page

Table 2.9 Continued from previous page

	(1) Model 1	(2) Model 2
	(0.072)	(0.073)
Stationary plant operators	1.000	1.036
	(0.066)	(0.069)
Assemblers	1.170**	1.178**
	(0.062)	(0.063)
Machine operators in agricultural/food industry, drivers & mobile plant operators	1.147***	1.204***
	(0.047)	(0.050)
Unskilled sales workers, cleaners and helpers	1.177***	1.109**
	(0.044)	(0.042)
Unskilled agricultural, forestry and fishery workers	1.053	1.002
	(0.069)	(0.066)
Labourers in mining, construction, manufacturing	1.219***	1.185**
	(0.066)	(0.064)
Education (ref: primary education)		
Secondary education		0.984
		(0.022)
Tertiary education		0.835***
		(0.033)
Marital status (ref: married)		
Widow		1.250***
		(0.046)
Separated/Divorced		1.387***
		(0.046)
Never married		1.587***
		(0.044)
Macro-region of residence (ref: Centre)		
Abroad		1.289*
		(0.139)
Islands		0.975
		(0.029)
North-East		1.085**
		(0.032)
North-West		1.164***
		(0.031)
South		1.022
		(0.025)
Social disability pension		8.477***
		(0.178)
Disability pension		1.413***
		(0.086)
Observations	361,829	361,829

Notes. Results from Cox proportional hazards models. Results are presented in the exponentiated form (hazard ratios). Standard errors in parentheses. *** p<0.001, ** p<0.01, * p<0.05, † p<0.1.

Table 2.10: Cox PH regressions - Women (macro-occupational groups)

	(1)	(2)
Occupation (ref: Clerical support workers)		
Managers and senior officials	1.000	1.211†
	(0.098)	(0.123)
Professionals	0.845***	0.987
	(0.037)	(0.048)
Technicians	0.948	1.000
	(0.045)	(0.048)
Service and sales workers	0.969	0.990
	(0.056)	(0.058)
Craft and related trade workers, skilled agricultural, forestry and fishery workers	0.950	0.919
	(0.069)	(0.068)
Plant and machine operators, assemblers	1.030	1.031
	(0.100)	(0.101)
Elementary occupations	1.050	0.981
	(0.054)	(0.052)
Marital status (ref: married)		
Widow		1.260***
		(0.050)
Separated/Divorced		1.513***
		(0.073)
Never married		1.762***
		(0.072)
Region of residence (ref: Centre)		
Abroad		1.462*
		(0.260)
Islands		0.917†
		(0.048)
North-East		1.193***
		(0.055)
North-West		1.225***
		(0.053)
South		0.972
		(0.042)
Education (ref: Primary)		
Secondary education		13.640***
		(0.494)
Tertiary education		1.673**
		(0.277)
Social disability pension		0.991
		(0.040)
Disability pension		0.917
		(0.053)
Observations	258,317	258,317

Results from Cox proportional hazards models. Results are presented in the exponentiated form (hazard ratios). Standard errors in parentheses. *** p<0.001, ** p<0.01, * p<0.05, † p<0.1.

Table 2.11: Cox PH regressions - Women (micro-occupational groups)

	(1) Model 1	(2) Model 2
Occupation (ref: General and keyboard clerks)		
Legislators and senior officials	0.908	1.099
	(0.114)	(0.143)
Managers	1.159	1.256
	(0.172)	(0.187)
Science, engineers, architects, life science, health professionals	0.754*	0.779*
	(0.090)	(0.099)
Legal, social and cultural professionals	1.002	0.989
	(0.065)	(0.066)
Teaching and research professionals	0.798***	0.951
	(0.041)	(0.054)
Science, engineering, life science technicians	0.957	0.964
	(0.064)	(0.065)
Business and administration technicians	1.018	0.981
	(0.069)	(0.067)
Public service technicians	0.864*	0.947
	(0.060)	(0.066)
Customer service clerks	0.974	0.746**
	(0.101)	(0.079)
Other clerical support workers	1.048	1.076
	(0.133)	(0.136)
Sales workers	0.964	1.009
	(0.090)	(0.095)
Personal service workers	0.986	0.946
	(0.095)	(0.092)
Personal care workers	0.922	0.847
	(0.127)	(0.118)
Protective service workers	0.975	0.966
	(0.092)	(0.093)
Mining, building and related trade workers	1.003	0.855
	(0.096)	(0.084)
Other craft and related trade workers	0.892	0.921
	(0.088)	(0.092)
Plant and machine operators	1.109	1.103
	(0.226)	(0.225)
Assemblers	1.011	0.971
	(0.110)	(0.106)
Unskilled sales workers	1.149*	1.029
	(0.073)	(0.066)
Cleaners and helpers	0.886	0.780*
	(0.090)	(0.084)
Unskilled workers in agriculture, forestry, fishery, mining, construction, manufacturing	0.966	0.898
	(0.086)	(0.080)
Education (ref: primary education)		
Secondary education		0.967
		(0.040)
Tertiary education		0.916
		(0.055)
Marital status (ref: married)		
Widow		1.258***

Continued on next page

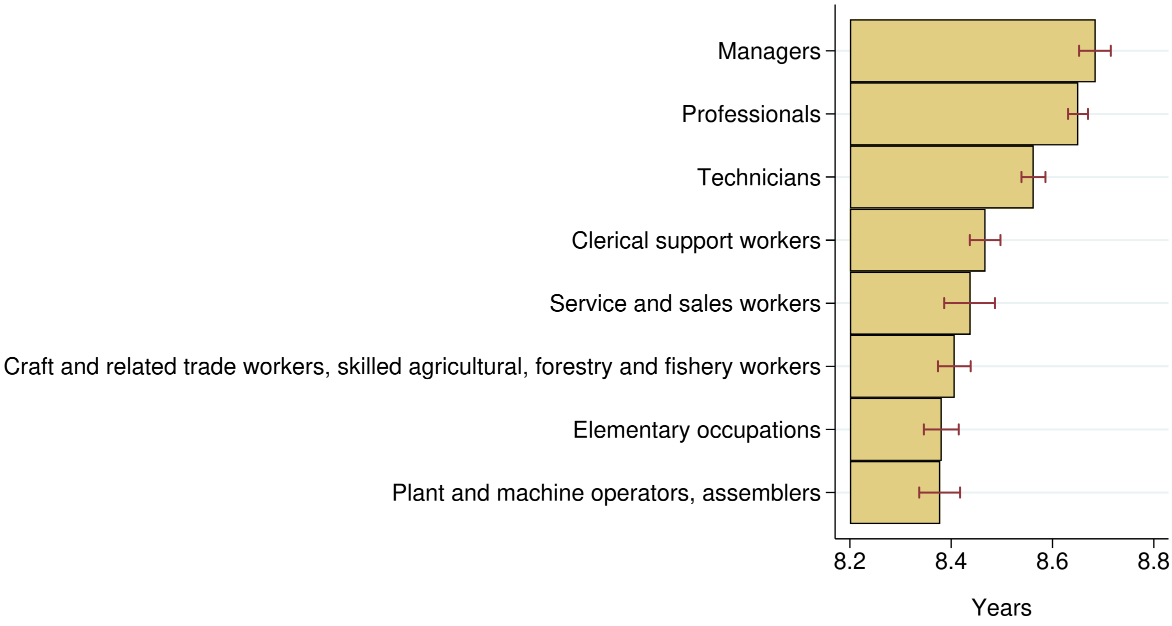
Table 2.11 Continued from previous page

	(1) Model 1	(2) Model 2
		(0.050)
Separated/Divorced		1.511***
		(0.073)
Never married		1.772***
		(0.073)
Macro-region of residence (ref: centre)		
Abroad		1.544*
		(0.278)
Islands		0.922
		(0.049)
North-East		1.195***
		(0.056)
North-West		1.228***
		(0.053)
South		0.977
		(0.043)
Social disability pension		13.659***
		(0.495)
Disability pension		1.696**
		(0.281)
Observations	258,317	258,317

Notes. Results from Cox proportional hazards models. Results are presented in the exponentiated form (hazard ratios). Standard errors in parentheses. *** p<0.001, ** p<0.01, * p<0.05, † p<0.1.

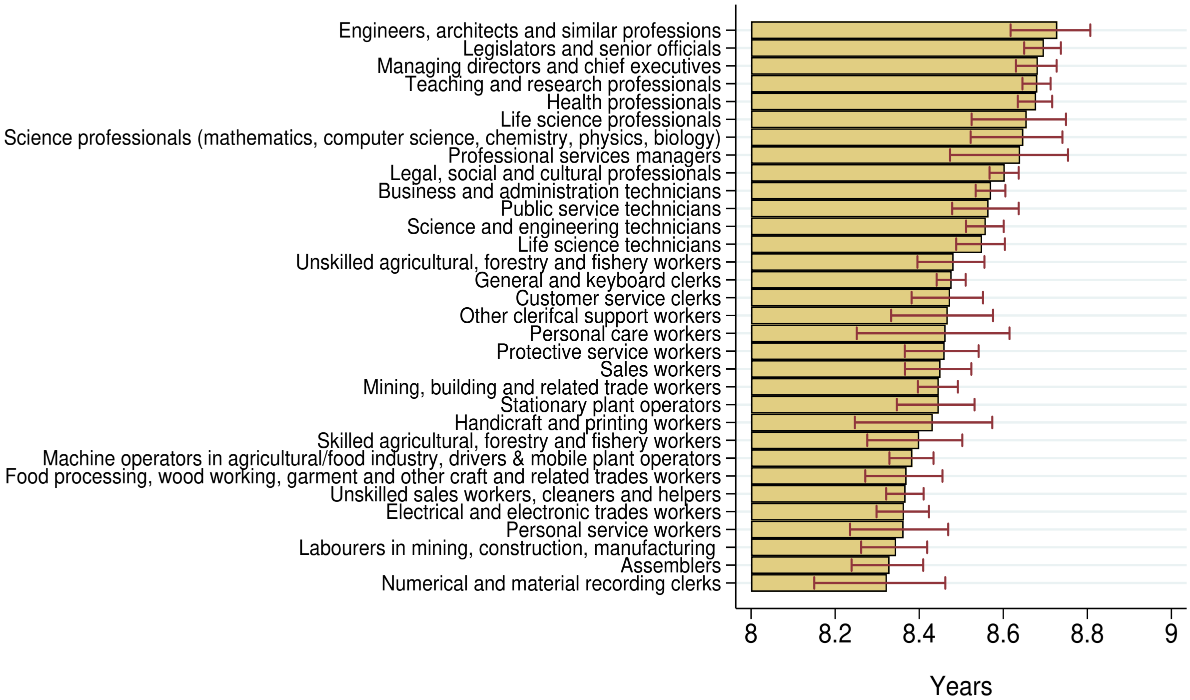
Figures Chapter 2

Figure 2.1: Partial life expectancy 65-74 (macro-class)
Men



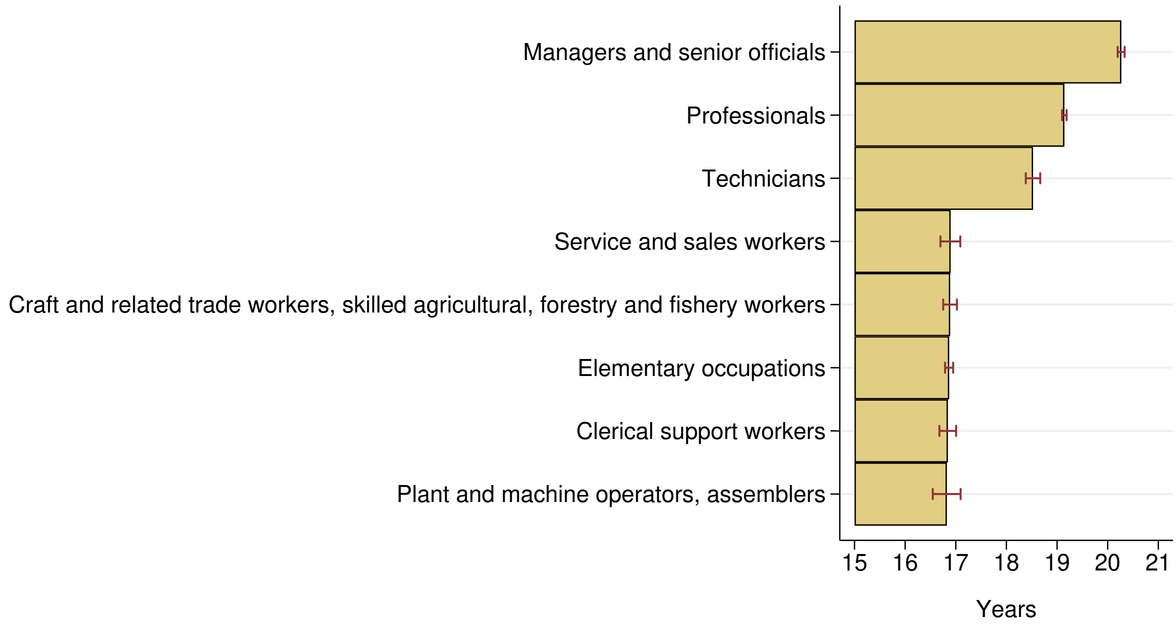
Notes. 95% confidence intervals in red.

Figure 2.2: Partial life expectancy 65-74 (micro-class)
Men



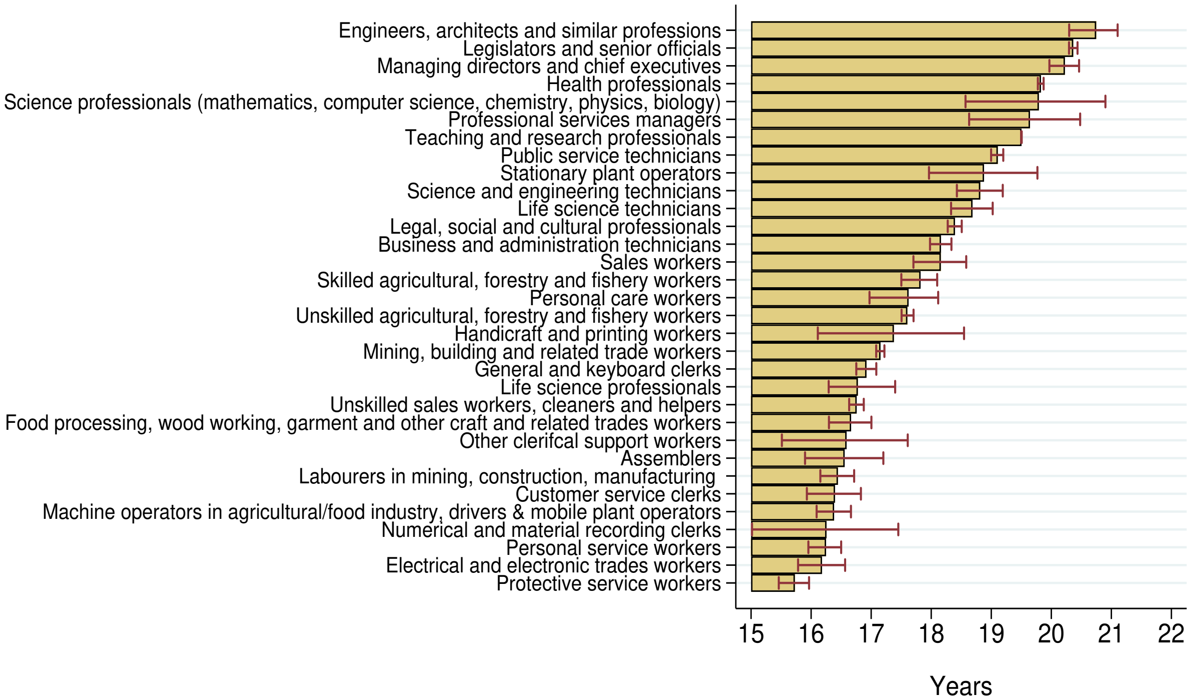
Notes. 95% confidence intervals in red.

Figure 2.3: Life expectancy at 65 (macro-class)
Men



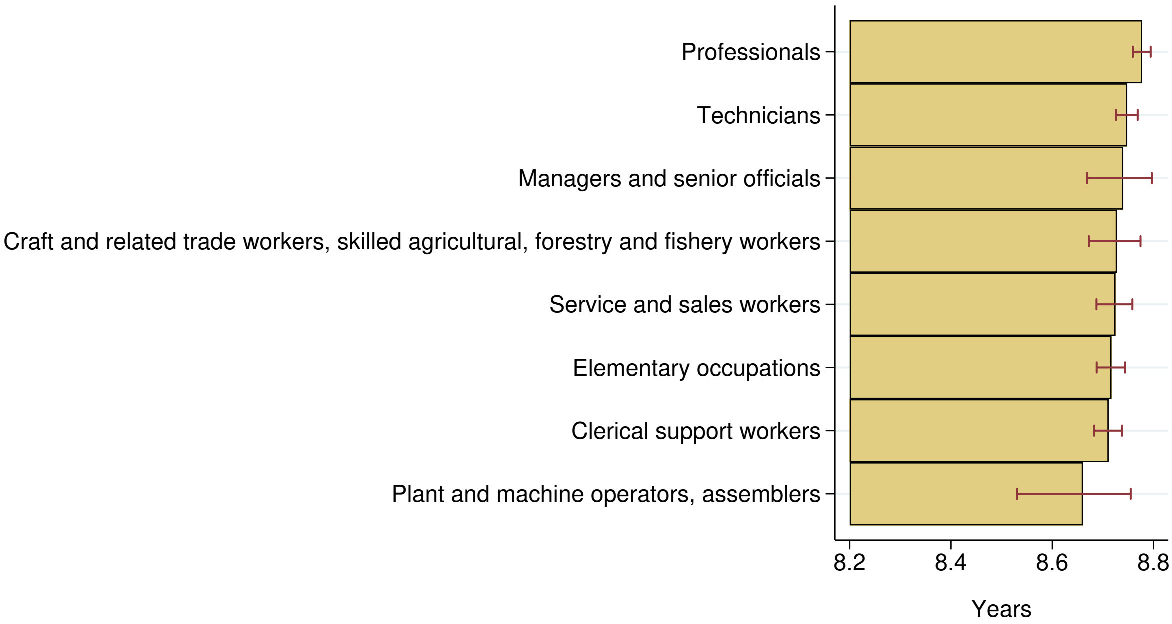
Notes. 95% confidence intervals in red.

Figure 2.4: Life expectancy at 65 (micro-class)
Men



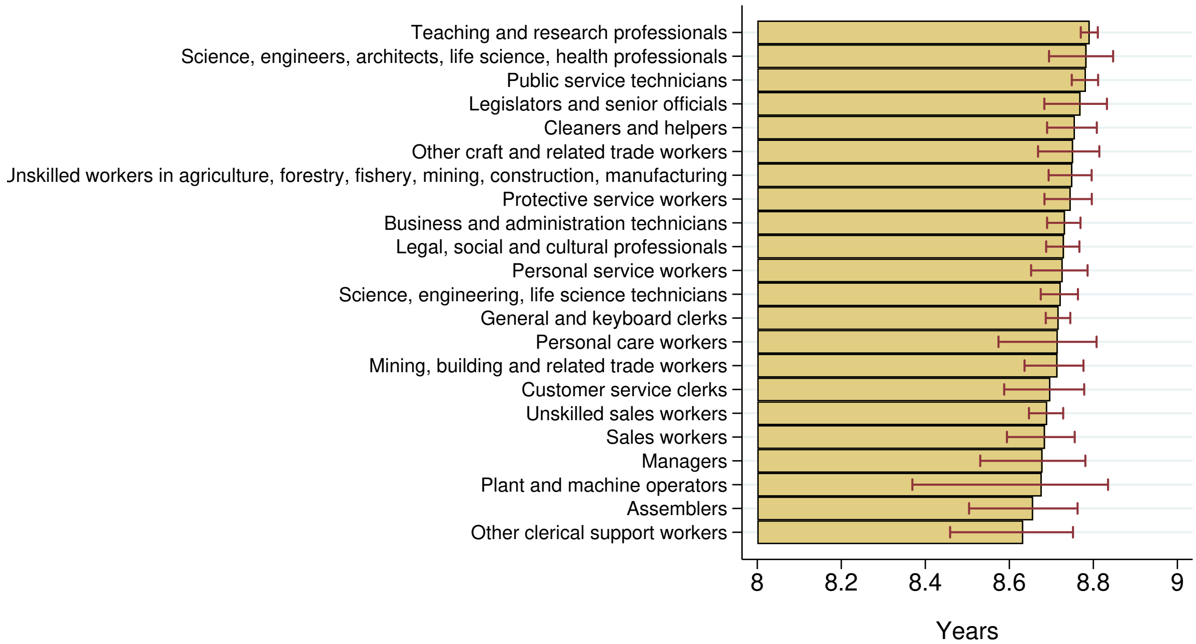
Notes. 95% confidence intervals in red.

Figure 2.5: Partial life expectancy 65-74 (macro-class)
Women



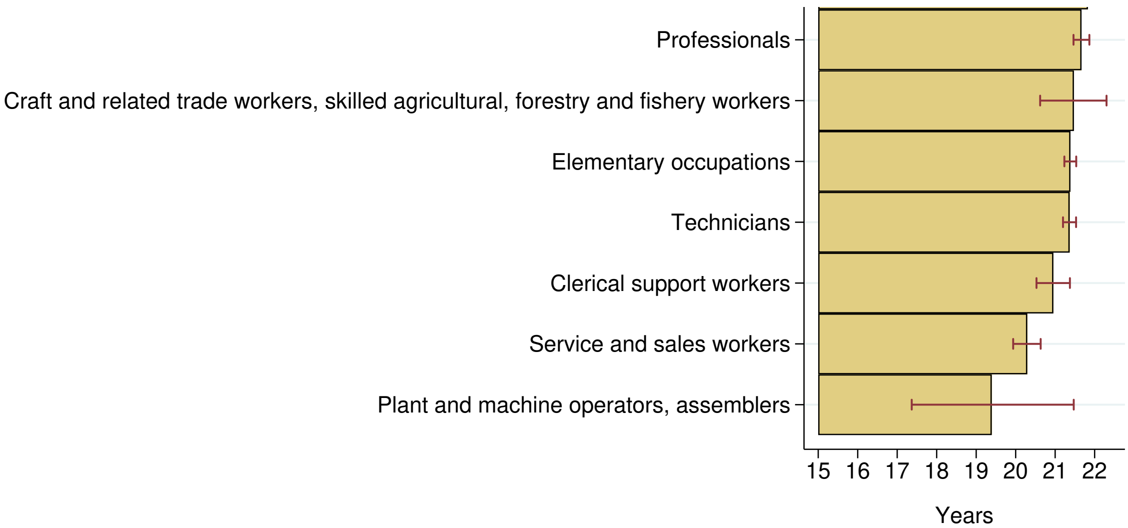
Notes. 95% confidence intervals in red.

Figure 2.6: Partial life expectancy 65-74 (micro-class)
Women



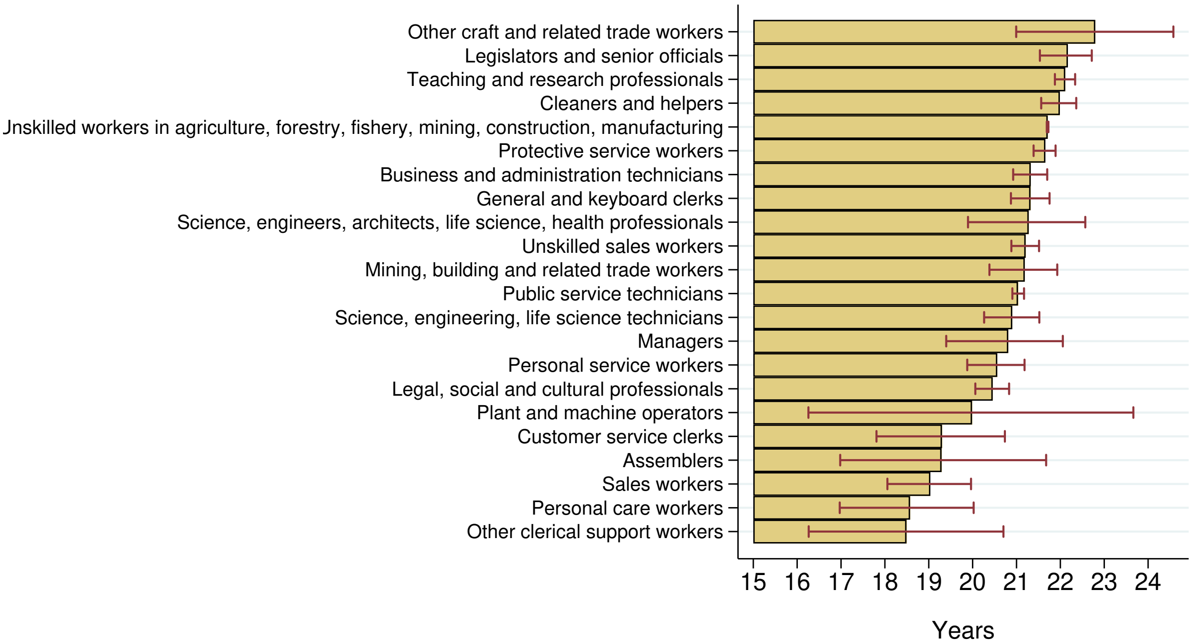
Notes. 95% confidence intervals in red.

Figure 2.7: Life expectancy at 65 (macro-class)
Women



Notes. 95% confidence intervals in red.

Figure 2.8: Life expectancy at 65 (micro-class)
Women



Notes. 95% confidence intervals in red.

Chapter 3

The evolution of lifespan inequalities among the over 50 in Italy and their distributional implications for the pension system

3.1 Introduction

In the last decades, rising inequality in socio-economic outcomes has emerged as a major theme in political and academic debates. There is mounting evidence that inequality in income, wealth, and lifetime earnings has been increasing over time and across cohorts in several OECD countries (Kopczuk et al. 2010, Piketty 2013, Saez & Zucman 2016, Bourgignon 2018). A key dimension of inequality, which is strongly related to economic inequality, is inequality in longevity (Cutler et al. 2006). While the causal nature of the relationship between socio-economic status and longevity remains a source of debate, the distribution of longevity is a key metrics of the distribution of well-being within a society. Moreover, lifespan inequalities have important implications when it comes to design and the evaluation of healthcare and social security programs (Auerbach et al. 2017). In this study, we provide novel evidence about lifespan inequalities, and about their evolution over time, in the Italian adult and elderly population, using multiple datasets extracted from the digital archives of the Italian Social Security Institute (INPS). We then evaluate

[1]This chapter is the result of a joint work with S. Ghislandi. We are thankful to F.C. Billari and to participants to the VisitINPS seminars for useful comments and suggestions. Access to administrative data was kindly granted by the Italian Social Security Institute as part of the VisitINPS Scholar Program. The opinions expressed in this study are those of the authors. They do not purport to reflect the opinions or views of the Italian Social Security Institute or its members.

the distributional implications of such inequalities for the Italian pension system.

Previous studies have documented the existence of tangible disparities in longevity by socio-economic status in Italy, especially among men (Lallo & Raitano 2018; Leombruni et al. 2015), discussing the challenges they pose to the equity and the sustainability of pension policies (Ardito et al. 2019; Caselli & Lipsi 2018; Mazzaferro et al. 2012). This literature presents, though, three main limitations. First, research on lifespan inequalities in the Italian population is overwhelmingly cross-sectional in nature. Analyses on the evolution of longevity differentials by socio-economic status are scant and yield mixed conclusions (Costa et al. 2017; Luy et al. 2015). This gap is particularly relevant in the light of the growing number of studies documenting widening inequalities in longevity by socio-economic status in several OECD countries (Case & Deaton 2021; Sasson 2016). Second, the literature on lifespan inequalities in the Italian context focuses exclusively on life expectancy differentials. As suggested by a growing strand of literature, though, life expectancy alone does not allow to fully capture the mortality profile of a given population, and of its evolution over time (Aburto et al. 2020). Specifically, life expectancy, which provides a concise and useful measure of average age-at-death, does not fully describe how deaths are distributed along the age distribution. In other words, it is not informative about the uncertainty surrounding the length of human life. Indeed, the relationship between 'central longevity indicators' (Cheung et al. 2005), such as life expectancy, and variation in age-at-death is not straightforward. In fact, while historically life expectancy and lifespan variation have been found to be inversely correlated, recent evidence suggests that such relationship may not hold for all countries or population subgroups (van Raalte et al. 2018). Monitoring both *average* age-at-death and *variation* in age-at-death is thus crucial for gaining full insight about heterogeneity in population health. Third, discussions about the distributional implications of differential mortality for the Italian pension system are not supported by evidence based on real employment and contributory histories. In fact, previous studies rely mostly on simulation exercises or back-of-the-envelope calculations to illustrate the regressive effects of unequal lifespans (Caselli & Lipsi 2018; Mazzaferro et al. 2012). As such, little is known about the real degree of redistribution implied by heterogeneity in longevity in the Italian pension system.

In this study we address all of these points. First of all, we document the evolution of life expectancy and lifespan variation by socio-economic status at 50 and at statutory retirement age among individuals with an employment background in the private sector in Italy. For this purpose, we rely on administrative data provided by the Italian Social Security Institute (INPS), encompassing the universe of employment spells in the Italian private sector over years 1975-2017 and the universe of pension benefits disbursed by INPS between 1995 and 2017. We use these data to build two measures of lifetime income, i.e. mid-career employment income and pension income, that we use as main markers of socio-economic status. We also use the available data to construct a measure of socio-economic status based on broadly defined occupational groups (blue-collar, white-collar, executives). Using mortality records spanning nearly four decades, we then show how lifespan inequalities have evolved over birth cohorts (1930-1957) and calendar years (1995-2017) for the chosen socioeconomic categories of interest. We further exploit the estimated cohort-specific mortality profiles to quantify the distributional implications of longevity differentials in the Italian pension system. Focusing on cohorts 1930-1950, we contrast the distribution of individual pension wealth and of the internal rate of return of pension contributions calculated under the assumption of homogeneous longevity with the distribution one obtains by accounting for heterogeneous mortality along the lifetime income dimension.

Our analyses confirm the existence of tangible disparities in life expectancy and document the existence of a clear socio-economic gradient also in terms of life span variation among men. Moreover, they suggest that, in the case of men, the socio-economic gradient in both life expectancy and lifespan variation has been worsening over the cohorts and the years under consideration. Instead, we find limited evidence of lifespan inequalities among women, detecting no significant trends over time. We also show that the distributional implications of unequal lifespans for the pension system are tangible. In particular, we document that the erosion in the profitability of pension contributions implied by heterogeneous longevity is stronger for male retirees at the bottom of the lifetime income distribution, and that such dynamics have become more pronounced over time.

The remainder of the study is organized as follows. In Section 3.2 we provide an overview of the main strands of the literature we contribute to. In Section 3.3 we present the datasets and the methodological approach we employ for deriving sex-specific mortality profiles differentiated by cohort/period and socio-economic status, together with the resulting estimates. In Section 3.4 we evaluate the distributional consequences of longevity differentials within the Italian pension system. Finally, Section 3.5 concludes, discussing the study results and limitations, together with venues for further research.

3.2 Background

This study integrates multiple strands of research on the evolution of lifespan inequalities by socio-economic status, and about their policy implications. A first strand of this literature has focused on the evolution of life expectancy differentials, particularly in the US. Recent works are unanimous in concluding that the longevity gap in the US has been rising over time, no matter how socio-economic status is measured. There is evidence of increase in life expectancy differentials by current income (Chetty et al. 2016), lifetime earnings (Waldron 2007, Cristia 2009, Bosworth et al. 2016, Auerbach et al. 2017), and education (Pijoan-Mas & Ríos-Rull 2014, Bosworth et al. 2016, Sasson 2016, Tan & Koedel

2019). In Europe, research has concentrated mostly on trends in life expectancy differentials by education (Murtin et al. 2017) and (broad) occupational class (Mackenbach et al. 2019). Research about evolution in the longevity gradient along the earnings distribution is less abundant, with some notable exceptions such as Denmark (Brønnum-Hansen & Baadsgaard 2012; Brønnum-Hansen 2017) and Germany (Kiebele et al. 2013, Wenau et al. 2019, Haan et al. 2020). For both countries, there is evidence of rising longevity gap by current income and lifetime earnings. A second strand of literature has sought to analyze the evolution of life expectancy differentials jointly with disparities in lifespan variation. As noted in the introductory section, lifespan variation encompasses a number of metrics capturing the dispersion of the age-at-death distribution. While life expectancy reflects the hypothetical average age-at-death in a population given its mortality profile, lifespan variation reflects the uncertainty surrounding such average. Although life expectancy and lifespan variation have been historically inversely correlated (Vaupel et al. 2011), a number of studies have shown that this relationship has been reversing in some countries or population subgroups, generally as a consequence of mid-life mortality crises which tend to display a clear socio-economic gradient (van Raalte et al. 2014; 2018). There is evidence of widening disparities in lifespan variation, due to lack of or slower compression of mortality among the most disadvantaged socio-economic groups, in Finland (van Raalte et al. 2014), in Denmark (Brønnum-Hansen 2017), in Spain (Permanyer et al. 2018) and in the US (Sasson 2016).

For what concerns Italy, researchers have documented tangible differences in life expectancy by education and occupational class, especially among men. Linking data from the 2011 Census with the Causes of Death Register (2012-2014), Petrelli et al. (2019) document a 3-year gap in life expectancy at birth between low and high educated men, and of 1.5 years in the case of women, in line with findings from previous works (Costa et al. 1994, Maccheroni 2008). Similarly, using social security data from INPS aggregated over years 1985-2012, Leombruni et al. (2015) find a gap of 1.8 years in residual life expectancy at 65 between former blue-collar workers and managers in the case of men. Lallo & Raitano (2018) combine social security data from INPS with survey data from the Italian 2005 EU-SILC module to estimate life expectancy at 60 by a measure of socio-economic status that encompass former occupational class and household self-reported economic conditions. Their estimates, based on mortality follow-up spanning years 2005 through 2009, imply a difference of 5 years in remaining life expectancy at 60 between individuals with opposite socioeconomic statuses. Analyses on the evolution of longevity differentials by socio-economic status are scanter and yield mixed conclusions. Focusing on the population of Turin, in the North-East of Italy, Costa et al. (2017) document that the longevity gap between individuals with low (primary) education and high (university) education from the early 1970s through the early 2010s has remained fairly constant at about 4-5 years for both men and women. Luy et al. (2015) apply the orphanhood method using data from the multipurpose survey on 'Family, welfare institutions, and childhood conditions' conducted by the National Institute of Statistics (Istat) in the years 1998 and 2003 to examine nation-wide trends in life expectancy by education and occupational class over the 1980-1994 period. They observe an increase in the longevity gap by education and occupational status at age 30 for men and a decrease in the case of women, a pattern attributable to differences in smoking habits. Belloni et al. (2012) study the association between pension income, used as a proxy for lifetime income, and mortality risk after 65 among Italian male retirees over the 1980s and the 1990s. While not estimating longevity differentials directly, their analysis suggests that the socio-economic gradient in old age survival remained stable for Italian male retirees over the observed period, after accounting for regional differences. As for the evolution of disparities in lifespan variation, no evidence is available for Italy. Our twofold contribution is thus to provide evidence about changes in both life expectancy and lifespan variation by socio-economic status in Italy. To the best of our knowledge, we are also the first to examine trends in lifespan inequalities in Italy taking both a cohort and period approach. Analyzing mortality from both perspectives is particularly crucial for informing public policies, notably pension policies. Indeed, previous studies from different institutional contexts show that using period instead of cohort life expectancy in the calculation of pension benefits may create implicit subsidies which put the sustainability and fairness of public pension systems at risk (Ayuso et al. 2020; Belloni & Maccheroni 2013).

Finally, our work speaks also to the literature about the distributional implications of differential mortality for social security programs. Evidence from the US suggests that widening longevity differentials have been increasingly offsetting the progressivity built in the Social Security benefit formula (Bosworth et al. 2016; Tan & Koedel 2019). Auerbach et al. (2017) estimate that diverging trends in life expectancy will cause the gap between average lifetime programme benefits received by men in the highest and lowest lifetime earnings quintiles to widen by 130,000$ over cohorts born between 1930 and 1960. Research based on Germany, where pension benefits have a stronger contributory link compared to the US, finds heterogenous longevity makes the pension system regressive, and that regressivity has been sharpening across cohorts (Haan et al. 2020; Whitehouse & Zaidi 2008). In Italy, research efforts have concentrated on the distributional implications of differential mortality under notional defined contribution pension rules, which will fully apply to cohorts born after the mid-1970s. Mazzaferro et al. (2012) run micro-simulations to compare pension contributions' profitability under defined benefit (DB) and notional defined contribution (NDC) pension rules for a representative sample of the Italian population consisting of individuals born between 1975 and 2000. Using cohort-invariant mortality rates differentiated by education, they show that while NDC rules improve inter-generational fairness, as compared to DB rules, they also imply redistribution from low to high socio-economic status individuals as they fail to account for heterogeneity in survival. Caselli & Lipsi (2018) use education-specific cross-sectional mortality data from the Italian National Institute of Statistics (ISTAT) to evaluate redistribution patterns across education levels under the Italian NDC scheme. Their analysis confirms that regressive distributional dynamics along the educational dimension are sizeable, and that they tend to become more accentuated as retirement age increases. We integrate previous studies by analysing the distributional implications of sex-, cohort- and lifetime income-specific mortality profiles for a large sample of Italian retirees born between 1930 and 1950. Unlike previous works, which rely on simulation exercises, we use high quality administrative data which enables us to reconstruct real employment and contributory biographies.

3.3 The evolution of lifespan inequalities among the over 50 in Italy

3.3.1 Data

We rely on two main sources drawn from the INPS archives. First, we make use of annual data taken from the *Dichiarazioni UniEmens* archive, which covers the universe of private employment spells in Italy between 1975 and 2017. Second, we employ annual data from the *Casellario Pensioni* archive, which gives access to the universe of all types of pension benefits disbursed by pension schemes supervised by INPS between 1995 and 2017. A major drawback of INPS data is the paucity of information about personal characteristics. We notably lack information about education, family status, and other family background characteristics. Information about place (province) of residence and marital status is available for beneficiaries of pension benefits only. For the latter, we can also match across spouses for years 1995 through 2017. For each individual in the INPS archives, we have information about month and year of birth and, when applicable, death (updated to 31 December 2018).

3.3.1.1 The *Dichiarazioni UniEmens* archive

The *Dichiarazioni UniEmens* archive reports detailed information about the universe of private employment spells registered in Italy since 1975. Our observation period ends in 2017. For each spell, we have information about gross earnings, the number of days, weeks, and months worked, the broad occupational category (blue-collar, white-collar, middle-manager, manager) and the kind of job contract (full-time versus part-time). We also have information about periods of absence from work due to family/parental leaves, sickness/injury, temporary suspension/reduction of working activities covered by the Wage Guarantee Fund (*Cassa Integrazione Guadagni*). Our original dataset consists of 306,930,929 observations, relative to 21,966,659 individuals. We use the *Dichiarazioni UniEmens* data to classify individuals by socio-economic status along two dimensions: employment income and broad occupational category, both measured at mid-career time (ages 45-49). As far as income is concerned, for each individual we calculate the inflation-adjusted average of non-zero gross employment earnings between ages 45-49, which we take as a proxy for lifetime income following the approach of previous studies (Auerbach et al. 2017; Bosworth et al. 2016; Milligan & Schirle 2021).[1] It is important to stress that our data do not allow to distinguish periods of missing earnings due to inactivity or unemployment, self-employment, work under social security schemes not managed by INPS (e.g. public employment), or informal work. By relying solely on calendar years with positive earnings, we intend to construct a measure of earnings potential (lifetime income) which is not affected by unemployment, severe health problems, or missing information on earnings from work not covered by our data. We then classify individuals into lifetime income quintiles based on the distribution of their birth cohort, considering men and women separately. As for occupation, we assign individuals to the prevalent occu-

[1] According to the economic literature on the inter-generational transmission of income inequality (Haider & Solon 2006), the age window which allows to minimize the life-cycle bias in the computation of lifetime earnings is lower than the one used in this study. For the Italian context, Bloise & Raitano (2021) find, for instance, that life-cycle bias is reduced if individuals are observed when they are around 35 years old. For the sake of comparison, though, we opted for sticking to the age window employed by previous works on mortality differentials by lifetime earnings.

pation category recorded between ages 45-49, distinguishing between blue-collar workers (*operai*), white-collar workers (*impiegati*) and managers (*dirigenti*).[2]

We restrict our analysis to individuals born between 1930 and 1957, for whom we can observe earnings and occupation between ages 45-49 and perform a mortality follow-up of at least ten years. Since we are interested in analyzing mortality after 50, we drop individuals who die before 50. Our final sample includes 6,949,246 individuals, 4,842,306 men and 2,106,940 women. Table 3.1 summarizes the steps taken in the construction of the sample, while Table 3.2 reports the distribution of the sample by cohort, together with survival information as of 2018. To get a sense of the evolution of lifetime income across cohorts, we plot trends in inflation-adjusted average permanent earnings by sex and year of birth in Figure C.1 (Appendix C). For men, average lifetime income increase steadily for cohorts born between 1930 and the mid-1940s. It stalls and then declines for later cohorts who were affected by the crisis of 1992 and by the recession of the early 2000s. Women, who record substantially lower lifetime income compared to men, experience a constant increase in average permanent earnings, except for later cohorts. As suggested by Figure C.2, which plots average permanent earnings by quintile and birth cohort, for men and women separately, the rise in average permanent earnings was driven mostly by individuals at the top of distribution, both for men and women. As for descriptive statistics by occupation, Tables C.2, C.3 and C.4 report the distribution of prevalent occupational position, by sex and year of birth, respectively. The most salient aspect is the scarcity of women among managers, especially in the case of earlier cohorts.

3.3.1.2 The *Casellario Pensioni* archive

The *Casellario Pensioni* dataset reports information about the universe of pension benefits disbursed by INPS between 1995 and 2017. For each pension benefit, we have

[2]The Italian labour law envisages a fourth broad occupational class consisting of white-collar employees with quasi-managerial responsibilities (*quadri* in Italian). Since this class is numerically residual and was introduced after 1985, we do not include it in our analysis.

information about the gross annual amount, the date in which the pension flow started, the date in which the pension flow ended (if this occurs by the end of 2017), the type of pension (old age, seniority, disability, etc.), the INPS pension scheme, and the years of contributions. We consider beneficiaries of pension benefits disbursed by INPS major pension scheme, i.e. FPLD (*Fondo Pensioni Lavoratori Dipendenti*), which represents the (public) pension scheme of private sector employees. We do not consider retirees covered by pension schemes for the self-employed, encompassing craftsmen, shop-keepers, and farmers, as for these categories pension represents a poor proxy of lifetime income (Belloni et al. 2012).[3] However, we consider former employees who also receive self-employment pension benefits if the latter represent a minor share ($< 50\%$) of their total pension income. We focus on beneficiaries of old-age (*vecchiaia*), seniority (*anzianità*) and early-retirement (*prepensionamento*) pension benefits, which are most clearly related to individuals' working life. We select individuals born between 1910 and 1950, who retired by and survive to age 67, which currently represents the statutory retirement age for calling old-age pension benefits. Since individuals belonging to these cohorts can be expected to retire at age 67 at the latest, age 67 is also the earliest age for which we can observe an entire distribution of pension benefits disbursed in each calendar year. We choose to focus on individuals born from 1910 onward because of the high proportion ($> 20\%$) of retirees with null or missing pension income among the older cohorts. We further select individuals who retired under defined-benefit or mixed (defined-benefit and defined-contribution) pension rules (further details below). Our final pensioners' sample includes 7,260,404 individuals 3,755,130 men and 3,505,274 women.

For each calendar year, we classify individuals into quintiles of pension benefits.[4] As-

[3]This is due to historically low contribution rates impinging on self-employed workers as compared to employees.

[4]The distribution of pension benefits in the case of women appearing in the *Casellario Pensioni* archive is strongly concentrated, particularly around minimum pension values. In order to assign women to pension quintiles, we add to each individual pension income a random amount between +5 and -5 euro.

signing individuals to pension benefit quintiles raises two main issues. First, to define quintiles consistently, one should ideally consider pension benefits drawn at a specific reference age, such as statutory retirement age. In our case, this can be done only partially, since retirees belonging to older cohorts can be observed for the first time well after statutory retirement age. We thus proceed as follows. In the case of individuals born from 1928 onward, we consider pension benefits collected at age 67. In the case of individuals born between 1910 and 1927, instead, we consider pension benefits collected at age attained in 1995 (our first year of observation), and use cut-off values from the pension benefits distribution at age 67 of cohort 1928. Second, pension rules changed repeatedly, and significantly, over the period of analysis. The first major change occurred with the Dini reform in 1995, which determined a slow transition from defined-benefit (DB) to notional defined-contribution (NDC) pension rules.[5] Indeed, the phase-in period of the Dini reform was set to be very gradual. Workers with at least 18 years of contributions as of December 1995 were fully unaffected by the reform. Instead, those with a shorter contributory record were to be affected on a *pro rata* basis, the weight of DB depending on the ratio between pre-1995 to the overall contribution period upon retirement. The second major change occurred in 2011, with the so called Fornero reform, which accelerated the transition to full NDC rules, introducing a *pro-rata* contribution for all workers starting from January 1, 2012. In other words, all pensions awarded from this date onward have an NDC component, regardless of the 18-year contribution period mentioned above. Over the period of analysis (1995-2017), retirees may thus belong to four main groups: (i) those who retired by 1995 fully under defined-benefit rules, (ii) those who retired between 1996 and 2011 fully under defined-benefit rules having accumulated at least 18 years of contributions as of December 1995, (iii) those who retired

[5]Under the defined benefit pension regime, pension benefits are determined multiplying pensionable earnings by the number of working years and by an accrual rate. Under the NDC regime, contributions are (fictitiously) accumulated in an individual fund, and are re-evaluated in line with a moving average of GDP growth. Pension benefits are then computed by multiplying the re-evaluated contributions by a coefficient which depends on remaining life expectancy at retirement. Such coefficients are neutral with respect to gender and other relevant socio-economic characteristics, but they are periodically updated to account for changes in official life expectancy projections.

between 1996 and 2011, under a mixed (*pro-rata*) regime having accumulated less than 18 years of contributions as of December 1995, (iv) those who retired between 2012 and 2017 under a mixed (*pro-rata*) regime. To account for differences in pension calculation formulas over the period of interest, we assign quintiles for those who retired fully under defined-benefit and mixed regime separately.[6] Table 3.3 shows the distribution by calendar year and pension regime, for men and women separately.

To gain insights about the difference implied by the two calculation mechanisms (defined-benefit vs mixed regime), in Appendix C we report the evolution of median inflation-adjusted annual gross pension amount at age 67 for cohorts born from 1928 onward, distinguishing pensioners who retired under defined-benefit rules from the rest of the sample. Figure C.3 plots trends for FPLD male and female pensioners, respectively. As one can notice, the difference implied by the calculation mechanism is particularly relevant in the case of women, who are clearly penalized by defined-contribution due to shorter/more fragmented careers and lower contributory amounts.

3.3.2 Methods

3.3.3 Period vs cohort mortality

We construct period- and cohort-based mortality profiles for different socio-economic groups depending on data availability. In the period-based approach, one considers the mortality experienced by individuals belonging to a given population (a 'synthetic' cohort) during a given period of time (e.g. a calendar year). In its most straightforward interpretation, period life expectancy measures the average lifespan that would prevail in the long run if the observed mortality conditions were held fixed. In the cohort-based approach, instead, one considers the mortality experienced by individuals belonging to

[6] We ignore individuals who retired fully under defined-contribution rules given their paucity during the period of interest.

the same birth cohort as the cohort ages. Cohort life expectancy thus measures the average lifespan that an individual belonging to a given cohort can be expected to live. While more suited to capture changes in mortality over time, the cohort-based approach is rarely adopted to track progresses in longevity. A major explanation for this lack of popularity owes to timeliness. Indeed, one needs a cohort to die out in order to compile its full mortality profile, which implies waiting at least a hundred years before being able to compute life expectancy for any given cohort (Guillot & Payne 2019). In this paper, when studying lifespan inequalities from a cohort perspective, we thus resort to projections based on observed, albeit incomplete, cohort mortality profiles.

3.3.3.1 Measuring lifespan variation

The demographic literature offers an array of indices of lifespan variation, which tend to be all highly correlated (van Raalte & Caswell 2013). In this paper, we opt for lifetable entropy, $\overline{H}$, which measures the elasticity of life expectancy with respect to mortality rates (Demetrius 1978; Keyfitz 1977; Keyfitz & Golini 1975; Leser 1955) and which has been used to study the evolution of lifespan inequality (Aburto et al. 2020). As shown by Goldman & Lord (1986) and Vaupel (1986), $\overline{H}$ can be expressed as follows:

$$\bar{H}(a) = \frac{e^{\dagger}(a)}{e(a)}$$

where $e^{\dagger}(a) = \int_{a}^{\infty} e(x)d(x)dx$, also known as 'life disparity', denotes the number of years lost to death at age a, and $e(a)$ denotes life expectancy at age a. As one can note, $e^{\dagger}$ is computed as the weighted average of the distribution of remaining life expectancies, $e(x)$, where the weights are given by the distribution of lifetable deaths $d(x)$. $\overline{H}$ is a dimensionless indicator, ranging between 0 and 1, where $\overline{H} = 1$ and $\overline{H} = 0$ indicate, respectively, maximum and null dispersion. As suggested by Alvarez et al. (2020), lifetable entropy is preferable to other indices of lifespan variation if one needs to compare lifespan variation at different ages. One reason is that lifetable entropy is, as said, a dimensionless indicator, as it does not depend on the level of mortality. In addition, unlike absolute measures of lifespan variation, such as life disparity or standard deviation, it does not hinge on the starting age of calculation a. Both properties are particularly useful if one needs to evaluate the distribution of lifespan at different starting ages, as in our case.[7]

3.3.3.2 Cohort life tables by mid-career income and occupational class

We construct life tables starting from age 50 by mid-career employment income and broad occupational group for cohorts born between 1930 and 1957, using data from the *Dichiarazioni UniEmens* dataset described in Section 3.3.1.1. We focus on individuals who survive to age 50, following them until 2018 or until the year of death, if this occurs earlier. Our data imply that we can construct heterogeneously incomplete mortality profiles across cohorts. Indeed, while we can observe individuals born in 1930 until age 88, for those born in 1957 the follow-up extends up to age 61 only. We tackle this issue by projecting mortality rates until 89 based on Gompertz' Law (Gompertz 1825), following the approach adopted by Chetty et al. (2016). The Gompertz' Law posits a log-linear relationship between mortality, M, and age, x, that is:

$$ln(M_x) = \beta_0 + \beta_1 x \tag{3.1}$$

Such relationship has proved to hold well to ages as old as 90 (Gavrilov & Gavrilova 2011). To build complete cohort mortality profiles we proceed as follows. First, we compute observed cohort mortality rates, by sex and socio-economic group, using data on deaths count and population exposure. Specifically, for cohorts born before 1943, we calculate observed mortality rates from 50 through 75, while for cohorts born from 1943 onward we calculate observed mortality rates from 50 through the last observable age:

[7] Despite its many useful properties, lifetable entropy is not readily amenable to within- and between-group decomposition. For this purpose, other kind of metrics of lifespan variability, such as the Theil index, could be considered (Permanyer et al. 2018).

$$M_{x,s,g,c} = \frac{D_{x,s,g,c}}{P_{x,s,g,c}}, \quad \forall x \in \{50, ..., a\} \tag{3.2}$$

where the subscripts x, s, g, c denote, respectively, age, sex, socio-economic group, and birth cohort, with $a = 75$ for cohorts 1930-1942 and $a = 2018 - c$ for cohorts 1943-1957. M denotes the mortality rate, computed as the ratio between deaths count, D, and population exposure, P, measured in terms of person-years lived. We use the observed mortality rates to estimate Gompertz' parameters β_0 and β_1 through ordinary least squares. For cohorts born before 1943, we estimate Gompertz' parameters between 50 and 75, while for cohorts born from 1943 onward we estimate Gompertz' parameters between 50 and the last observable age. We then use $\hat{\beta}_0$ and $\hat{\beta}_1$ to project mortality rates from age a to age 89, as $\widehat{M_{x,s,g,c}} = e^{\hat{\beta}_0 + \hat{\beta}_1 x}$. After 89, we apply age- and sex-specific mortality rates based on the Italian population taken from ISTAT referring to year 2018.

We build cohort life tables for each subgroup starting from mortality rates by applying standard lifetable techniques (Wachter 2014), from which we extract lifespan indicators of interest, i.e. life expectancy and lifetable entropy. To build confidence intervals for life expectancy and lifetable entropy estimates, we draw new Gompertz parameters from a multivariate normal distribution with the means and covariance matrix obtained directly from the ordinary least squares estimation procedure for each sex, socio-economic group and cohort combination (Chetty et al. 2016). We make 1,000 draws for each sex, socio-economic group and cohort combination.[8] We then form our confidence intervals using the 2.5th and 97.5th percentiles of the resulting life expectancy and lifetable entropy distribution for each combination of interest.

[8]In other words, we perturb estimated Gompertz coefficients in order to build 1,000 life tables for each sex, socio-economic group and cohort combination.

3.3.3.3 Period life tables by pension income

We construct period life tables by pension income quintile starting from statutory retirement age, currently set at age 67, for calendar years 1995-2017, using data from the *Casellario Pensioni* archive, described in Section 3.3.1.2. As said, we keep beneficiaries of old-age (*vecchiaia*), seniority (*anzianità*) and early (*pre-pensionamento*) pension benefits, born between 1910 and 1950, who retired by 67. After such age, we can expect most individuals to be retired. Our data allow us to follow the selected individuals from 1995 to 2017 or until the year of death, if this occurs earlier. This implies that we can track mortality from age 67 to age attained in 2017 for those born from 1928 onward, and from age attained in 1995 to age attained in 2017 for those born prior to 1928. The nature of these data makes cohort-based analysis poorly suitable for studying the evolution of lifespan inequalities at pensionable age. Indeed, for younger cohorts, i.e. those born from 1928 onward, the age span over which one may estimate reliable Gompertz' parameters for further projections is too limited, if not null, while for older cohorts, i.e. those born prior to 1928, we are unable to observe mortality at 67. For these reasons, we adopt a period-based approach, implementing the following steps. First, for ages 67 through 85, we compute sex-, age-, and pension income quintile-specific mortality rates for each calendar year, as the ratio between deaths count and population exposure (person-years lived):

$$M_{x,s,q,t} = \frac{D_{x,s,q,t}}{P_{x,s,q,t}}, \quad \forall x \in \{67, ..., 85\} \tag{3.3}$$

where the subscripts x, s, q, t denote, respectively, age, sex, pension income quintile, and calendar year. We compute mortality rates based on observed data until 85 only, as this is the oldest age for which mortality is observable in 1995. We then use the observed mortality rates to extrapolate mortality rates at older ages (85+) by applying the Kannisto model, which is best suited for approximating mortality at very old ages (Thatcher et al.

1998). According to the Kannisto model, mortality at older ages can be approximated as follows:

$$M_x = \frac{\alpha e^{\beta x}}{1 + \alpha e^{\beta x}} \tag{3.4}$$

We fit the Kannisto model through ages 75-85 and estimate parameters α and β through maximum likelihood for each sex, quintile and calendar year combination (see Appendix C.1), and use the estimated Kannisto parameters to extrapolate mortality rates from 86 to 120. We build sex-specific period life tables for each pension quintile and calendar year starting from mortality rates. We estimate confidence intervals for life expectancy and lifetable entropy estimates by bootstrapping using Monte Carlo simulation methods, assuming deaths count follow a binomial distribution (Andreev & Shkolnikov 2010; Chiang 1984). Since the *Casellario Pensioni* dataset reports also information about pension beneficiaries' place (province) of residence, we construct sex-specific period life tables by pension quintiles both at the national and at the regional level. When constructing regional-level life tables, we assign individuals to pension quintiles based on region-specific pension distribution.

3.3.4 Lifespan inequalities at 50

3.3.4.1 Main results

In this section, we show how lifespan inequalities at 50 by socio-economic status have evolved over cohorts 1930-1957. Figures 3.1 and 3.2 display the evolution of life expectancy at 50, by lifetime employment income and occupational group, for men and women separately. Point estimates for each sex and socio-economic group have been averaged by decade of birth (single-cohort life tables are reported in Appendix C). Results by lifetime income yield two main findings. First, there is a clear gradient in life expectancy in the case of men, which turns to be widening across cohorts, notably for individuals at the bottom of the distribution. Indeed, the gap in life expectancy at 50 between top and bottom quintile rises from approximately from 3 to 4.5 years comparing men born in the 1930s and in the 1950s. For women, on the contrary, we do not observe a clear gradient in life expectancy across the lifetime income distribution. Differences across quintiles are minimal and statistically not significant, as shown by overlapping confidence intervals. Results by broad occupational class yield a slightly different picture. We document a steep gradient in the case of men, which is stable across selected cohorts. Specifically, white-collar and managers boast an advantage of about 2 and 4.5 years respectively as compared to blue-collar workers. A modest gradient appears also in the case of women. Still, estimates come with considerable uncertainty, as suggested by the width of confidence intervals. Figures 3.3 and 3.4 display the evolution of lifetable entropy at 50, by lifetime employment income and occupational group. In the case of men, we observe a clear gradient, which appears to be widening across cohorts, along both lifetime income and occupational group. In the case of women, we find a small gradient when measuring socio-economic status by occupational group, while differences across the income distribution remain negligible.

3.3.4.2 Sensitivity analysis

Our estimates of life expectancy and lifetable entropy at age 50 by lifetime employment income and occupation depend on Gompertz parameters which, for cohorts born from 1943 onward, are estimated over different age ranges. It is thus possible that widening inequalities observed by income in the case of men are due to selection effects rather than by truly worsening survival disparities. We address this issue in a twofold way. First, we calculate survival probabilities over ages 50-61, as this is the age range over which we can observe mortality for all the selected cohorts. Results, reported in Appendix C, are in line with those presented in the previous section. As one can see from Figure C.6, for individuals born in the 1930s, the difference in the probability of surviving to 61 at age

50 at the bottom and at the top of the permanent earnings distribution is about 2.5pp. For individuals born in the 1950s, the difference raises to 3.5pp. Likewise, Figure C.7 confirms the stability of survival differentials along the occupational dimension. Second, we construct cohort life tables based on observed mortality rates only. Specifically, for all cohorts born after 1928, we replace unobserved mortality rates until age 89 with those from the closest cohort for which such mortality rates are observed. For instance, we replace unobserved mortality rate at age 89 of cohort 1929 with that observed for cohort 1929. Similarly, we replace unobserved mortality rate at ages 88 and 89 of cohort 1930 with those observed for cohorts 1928 and 1929 respectively, etc. Also in this case, results are qualitatively in line with those discussed in Section 3.3.4.1, both for mid-career earnings and occupational position (see Figure C.8 and Figure C.9).

Another issue arises from the presence of individuals with relatively low attachment to the labour market, at least in terms of private employment. Indeed, about 23% of individuals in the *Dichiarazioni Uniemens* dataset are observed for less than 4 years between ages 45-49 (Table C.1). These might be individuals who alternate periods of dependent and autonomous work, or who transition back and forth between formal and informal work, or move into or out of the public sector. These might also be individuals who leave the labour market for family or health reasons or who migrate abroad. In all these cases, average employment income between ages 45-49 measurable through the *Dichiarazioni Uniemens* dataset might be a poor proxy of permanent earnings. It is possible, for instance, that people in the bottom mid-career income quintile are individuals with low labour market attachment due to poor health or family issues. To attenuate these concerns, we repeat our analysis by dropping individuals who are observed for less than 4 years between 45 to 49 in the *Dichiarazioni Unienems* dataset. This leaves us with a sample of 5,373,049 individuals, 3,876,087 men and 1,496,962 women. Results, presented in Appendix C, are analogous to baseline ones, for both men and women (Figure C.10 and Figure C.11).

The same holds when performing such sensitivity analysis on results relative to prevalent mid-career occupational position (Figure C.12 and Figure C.13).

3.3.5 Lifespan inequalities at statutory retirement age

3.3.5.1 Main results

Figure 3.5 and Figure 3.6 display the evolution of life expectancy and lifetable entropy at 67 by pension income at the national level for former private employees (FPLD), over calendar years 1995-2017, grouped by 5-year periods (3-year in the case of 2015-2017; estimates by single calendar year are reported in Table C.9 to Table C.12). We observe a widening gradient both in life expectancy and lifetable entropy by pension quintile among men. Over the examined period, the gap in LE at 67 between male FPLD pensioners at the top and the bottom of the pension income distribution doubles. Instead, we observe no such a gradient among women. If anything, female FPLD pensioners at the bottom of the pension income distribution appear to fare better in terms of both average and variation in age-at-death compared to counterparts at the top of the pension income distribution.

We exploit the information about pensioners' province of residence to check whether such trends apply to the whole country. Figure 3.7 and Figure 3.8 plot the difference in life expectancy at 67 between individuals located at the top and at the bottom of the pension income distribution, for men and women separately, across Italian regions, in 1995-1999 and in 2015-2017. A few results stand out. First, among men, the life expectancy gap is larger among Northern than Southern-Central regions, with the notable exceptions of Molise.[9] It is noteworthy, though, that the life expectancy gap among men has widened in most Italian regions when comparing 1995-1999 to 2015-2017. In the case

[9]Northern regions include Valle d'Aosta, Lombardia, Piemonte, Liguria, Veneto, Trentino Alto-Adige, Friuli-Venezia Giulia, Emilia-Romagna; Central regions include Toscana, Lazio, Umbria and Marche; Southern regions (including Islands) include Abruzzo, Molise, Puglia, Campania, Basilicata, Calabria, Sicilia and Sardegna.

of women, region-level analysis confirms that differences in life expectancy at the top and at the bottom at the income distribution among women are mostly negligible (± 1 year), and that in some regions women at the bottom of the pension income distribution can expect to live *longer* than women at the top of the income distribution. Although some regions display some visible changes when comparing life expectancy differences in 1995-1999 against 2015-2017 (e.g. Marche and Valle d'Aosta), we detect no clear trend in the evolution average lifespan disparities among female retirees across Italian regions, consistently with nation-level results.

3.3.5.2 Sensitivity analysis

A possible explanation for the lack of a clear gradient in both life expectancy and lifespan variation among female FPLD retirees is that pension income is a poor proxy of socio-economic status for this category of women. It is possible, for instance, that women at the bottom of the pension income distribution are women who could afford to hold low-paying jobs or to opt for fragmented careers based on the availability of household resources.

A first way to test for this hypothesis is to restrict the analysis to women with pension income above the minimum in each calendar year[10], i.e women whose pension income could be more revealing of socio-economic status. Still, (nation-level) results reported in Figures C.14 and C.15 in Appendix C are qualitatively and quantitatively analogous to those presented in the previous section, with no statistically discernible socio-economic gradient in both life expectancy and lifespan variation among female retirees.

An alternative approach is to relate women's post-retirement mortality to husbands' pension income, which serves as a proxy for household resources. For this purpose, we exploit

[10] The pension minimum for calendar years 1995-2017 is reported in Table C.14.

the possibility of matching across spouses appearing in the *Casellario Pensioni* archive. Since the latter keep tracks of pension benefits disbursed between 1995 and 2017, we can only focus on female FPLD pensioners married to FPLD beneficiaries of old-age/seniority pension benefits who were still alive in 1995. We restrict our analysis to women turning 67 between 1995 and 2017, i.e. those born from 1928 onward. We perform a mortality follow-up which extends up to 2017, or until their year of death, if this occurs earlier. We apply logistic survival analysis where yearly mortality risk of female FPLD pensioners is regressed against husband (cohort-specific) pension quintile, own (cohort-specific) pension quintile, year of birth, age difference with respect to husband, widowhood status, macro-region of residence, and macro-region of birth. We opt for not imposing any constraint on the baseline hazard. That is, we include in the model as many dummies as the maximum survival time observed. Since the earliest and last year of observation are 1995 and 2017 respectively, the maximum survival time is 23 years. Table 3.4 reports the results, in terms of odds ratios. Two findings stand out. First, in line with results presented in Section 3.3.5.1, we find that women's post-retirement mortality correlates *positively*, albeit non linearly, with their own pension quintile. Women in the top pension quintile display a post-retirement mortality risk which is 16.5% higher than women in the bottom pension quintile, *ceteris paribus*. Second, we document the existence of an inverse relationship between women's post-retirement mortality and husband's pension income, which appears in particular at the top of the distribution. Women whose husbands locate in the respective top pension quintile have a post-retirement mortality risk which is 12% lower than women whose husbands locate at the bottom of the male pension income distribution, *ceteris paribus*. While data do not allow to properly study lifespan inequalities by spouse's income, these findings question the limited, if not inverse, socio-economic gradient in health and longevity among elderly women documented in the previous section, as well as in previous works (Leombruni et al. 2015), based on women's own income and occupation.[11]

[11]We also analyze association between men's port-retirement survival and wife's pension income. We find a negative association between men's mortality and wives' pension quintile, but for men whose wives

3.4 The distributional implications of longevity differentials

In this section, we evaluate the distributional consequences of longevity differentials within the Italian pension system. We focus on individuals born between 1930 and 1950 who retired under the private employees' pension scheme (FPLD). The frequency of changes in pension rules over the last decades in Italy, briefly outlined in Appendix C, poses some challenges. In particular, it makes difficult to disentangle changes in distributional dynamics due to changes in longevity differentials across birth cohorts, as individuals may face different retirement conditions depending on the retirement timing. While most individuals in cohorts 1930-1950 had their pension computed under defined-benefit rules, policy changes touching upon the contributory/age requirements, the reference period for the calculating pensionable earnings and the indexation mechanism, may imply that important discontinuities in retirement conditions exist within and between birth cohorts. With this caveat in mind, we proceed as follows. We first reconstruct the contributory biographies of a large sample of FPLD retirees based on the *Estratti Conto* INPS archive. The *Estratti Conto* archive allows to track the contributory history of private sector workers covered by INPS-managed social security schemes. The *Estratti Conto* dataset provides a detailed record of all episodes in one's working life covered by INPS social security contributions: employment/self-employment job spells, parental/family leaves, sickness/injury episodes, unemployment spells covered by social benefits. Thus, compared to the *Dichiarazioni Uniemens*, the *Estratti Conto* allow to track spells falling outside the scope of private employment episodes. We then relate contributory histories to observed/estimated pension flows in order to evaluate distributional dynamics across the chosen cohorts. We look at two main measures: pension wealth at retirement (PW) and the internal rate of return of pension contributions (IRR). To isolate the distributional implications of longevity differentials, we compare the distribution

locate at the top of the female pension distribution (Table C.13).

of PW and IRR based on cohort-, sex- and lifetime income quintile-specific mortality profiles previously estimated, with the distribution of PW and IRR based on cohort- and sex-specific mortality profiles, averaged across all lifetime income quintiles.

3.4.1 Data and methods

3.4.1.1 Pension contributions

We have access to the contributory histories, reported in the *Estratti Conto* archive, of a random sample of 260,584 FPLD retirees, who appear as recipients of old age, seniority or early retirement pension benefits in any year between 1995 and 2017 in the *Casellario Pensioni* archive.[12] Out of this initial sample, we select individuals born between 1930 and 1950 who could be assigned to cohort-specific lifetime income quintile based on average gross employment earnings observed at ages 45-49 as recorded in *Dichiarazioni Uniemens* (Section 3.3.1.1). We further focus on the subset of "stable" workers, i.e. individuals who whose contributory history, as tracked by the *Estratti Conto* dataset, does not contain substantial gaps (>5 years). We end with a final sample composed of 97,321 individuals (69,241 men and 28,180 women). Table 3.5 recapitulates the steps taken in the construction of the sample.

The *Estratti Conto* archive presents two major limitations for our purposes. First, for each contributory spell it does not report the amount of contributions, but only the social security taxable base, i.e. gross earnings. For this reason, we focus on contributory spells falling under the FPLD fund, as for the latter reliable information on historical contributory rates is made directly available by INPS (Figure C.4). Second, the *Estratti Conto* archive provides reliable information about the social security taxable base from 1974 onward. Prior to 1974, information about gross earnings is missing for most contributory spells. This means that for most cohorts we have only a partial overview of

[12] Individuals were randomly selected on the basis of their unique identifier, conditional on appearing in the *Casellario Pensioni* as recipients of old age, seniority and early retirement pension benefits.

individual lifetime earnings profiles, and consequently of their contributory biographies. In order to construct a measure of contributions paid over one's entire working life, which is crucial for our distributional analysis, we need first of all to impute gross earnings for years prior to 1974. This implies we need also to impute the age at which individuals start working for those who enter the labour market prior to 1974. For this purpose, we exploit information about years of contributions available for each FPLD retiree. Following Brugiavini & Peracchi (2003), we impute gross annual earnings for years prior to 1974 by fitting a simple fixed effects model for the logarithm of gross earnings, using age and the years of contribution as predictors, where age enters as a cubic polynomial while years of contribution enter linearly. The model looks as follows:

$$log(\text{Earnings})_{it} = \beta_0 + \beta_1 \text{Age}_{it} + \beta_2 \text{Age}^2_{it} + \beta_3 \text{Age}^3_{it} + \beta_4 \text{Years of contributions}_{it} + \gamma_i + \epsilon_{it} \tag{3.5}$$

where subscripts i and t identify individuals and years, respectively, and γ_i identifies individual fixed effects. We fit the model separately for men and women through ages 24-60 focusing on full-year, inflation-adjusted earnings.[13] Table 3.6 reports the estimated regression coefficients, that we use to impute earnings for years prior to 1974. Figure C.5 plots the life cycle earnings profile predicted by the model, for men and women separately. The implied growth rate of full-year earnings declines with age, the earnings level flattening around 50. Finally, for each retiree, we estimate the amount of individual contributions in each year by multiplying the imputed/observed annual gross earnings by the relevant contribution rate.

[13]That is, for each observation we exclude first and last year in the sample to account for the fact that people typically work only part of the year they enter / exit the sample. The model is estimated through ages 24-60 as 24 is the first age observed in the Estratti Conto for individuals born in 1930-1950, and average retirement age for individuals in the sample is below 60.

3.4.1.2 Pension wealth

Pension wealth at retirement (PW) can be expressed as the sum of pension benefits received from retirement R until the expected age of death T:

$$PW_i = \sum_{s=R_i}^{T_i} p_{is} \tag{3.6}$$

where p_{is} represents pension benefits received by individual i at age s, R is the age at retirement, and T is the expected age of death at retirement. Since the period coverage of *Casellario Pensioni* starts in 1995, we calculate pension wealth of individuals who survive to age 65, as this is the first age at which individuals born in 1930 can be observed in the dataset. We are able to calculate pension wealth at retirement for 93,260 individuals (Table 3.5). Since we have information about gross annual pension benefits from 1995 up to 2017, and there are some individuals who retire before 1995 and are still alive at the end of 2017, we must estimate the stream of individual pension benefits before 1995 and beyond 2017. For years prior to 1995, we discount annual pension benefits received in 1995 by the average growth rate of pension benefits observed from 1995 onward for each individual. For years between 2018 and 2021, we let annual pensions to grow following the official indexation schedule provided by INPS.[14] After 2021, we assume that the same indexation schedule of 2021 holds, with an inflation rate of 1.5%. Finally, we discount compound/discount all pension benefits to the year 2019 with the consumer price index/projected inflation.

[14]Pension benefits in Italy are indexed to price inflation since 1993. The indexation rate is not homogeneous, but it depends on pension amount. Pensions below a certain threshold, defined as three times the minimum pension, are indexed perfectly. Pensions above the threshold are indexed only partially, at a decreasing rate. For each individual, INPS considers the relevant indexation rate based on the sum of all pension benefits, including disability pension benefits, survival pension benefits, etc. Here, we make the simplifying assumption that individuals receive old-age, seniority and early retirement pension benefits only.

3.4.1.3 Internal rate of return

From an individual perspective, pension wealth *per se* is not informative about the profitability of pension contributions. For this purpose, one needs a measure which allows to compare lifetime contributions with expected pension benefits. Here, we opt for the internal rate of return (IRR). The IRR can be seen as the interest rate that should apply on contributions for generating enough funds to ensure the pension flow from retirement until the (statistical) death of the individual. In practice, the IRR is the interest rate r that equalizes the stream of pension contributions c paid over one's working life and the expected stream of pension benefits p at retirement R:

$$\sum_{t=a_i}^{R_i-1} c_{it}(1+r_i)^{R_i-t} = \sum_{s=R_i}^{T_i} \frac{p_{is}}{(1+r_i)^{s-R_i}} \tag{3.7}$$

where c_{it} is the contribution paid by individual i at age t, p_{is} is the pension benefit received at age s, r_i is the internal rate of return, a_i is the age at entry in the labour market, R_i is the age at retirement, and T_i is the expected age of death at retirement. An alternative to IRR is represented by the net present value ratio (NPVR), defined as the ratio between the present value of pension benefits and the present value of lifetime contributions (Mazzaferro 2019; Mazzaferro et al. 2012). Albeit computationally simple, a major limitation of the NPVR is that it requires some arbitrary choice for the discount rate. Instead, the IRR allows to endogenously determine the rate of return which equalizes the present value of pension benefits and the present value of contributions. We calculate individual rates of return under heterogeneous and homogeneous longevity using contributions and observed/estimated pension benefits derived as described above. We compound/discount all contributions and pension benefits to the year 2019 with the consumer price index/projected inflation.

3.4.2 Results

Figure 3.9 and Figure 3.10 plot the distribution of pension wealth by lifetime income quintile across birth cohorts calculated with heterogeneous and homogeneous life expectancies, for men and women respectively. In both scenarios, pension wealth increases monotonically with lifetime income quintile as pension benefits are computed over higher lifetime earnings, in real terms. Distributional implications of heterogeneous longevity are strongly gender-specific. In the case of men, we find that pension wealth is higher under the assumption of heterogeneous longevity for the top two quintiles, while it declines for the lower quintiles. Consistent with widening longevity differentials, such patterns become more pronounced across cohorts. As shown by Figure 3.11, the top quintile of male cohorts 1930-1934 gains about 58,000€ (+5.3%) in pension wealth when accounting for heterogeneous mortality, while the bottom quintile loses about 13,000€ (-5.7%). For male cohorts born in 1945-1950, the gain and the loss implied by heterogenous longevity amount to about 110,000€ (+10.8%) and 38,000€ (-10.1%) for the top and bottom quintile, respectively. Trends in cohort-specific Gini coefficients of pension wealth among men corroborate these results (Table 3.7, Panel A). While the Gini coefficients of pension wealth shrinks over the selected male cohorts, the difference between Gini coefficients under heterogenous and homogeneous longevity increases from 0.0192 to 0.0292 between male cohorts 1930-1934 and 1945-1950. In the case of women, distributional consequences of heterogeneous longevity are hardly sizeable, and do not follow a clear pattern across lifetime income quintiles and cohorts (Figure 3.12). These findings are consistent with disparate trends in life expectancy by lifetime income among women documented in the previous part of the paper. Similarly, differences in Gini coefficients under heterogeneous and homogeneous longevity are also minimal in the case of women (Table 3.7, Panel B). We get analogous results when investigating trends in pension wealth by retirement year rather than by birth cohort.

Figure 3.13 and Figure 3.14 shows the distribution of IRR (%) by lifetime income quintile across birth cohorts calculated with heterogeneous and homogeneous life expectancies, for men and women respectively. A few patterns emerge. First of all, average internal rates of return for each sex-quintile-cohort combination are all largely positive, both under the heterogeneous and homogeneous longevity scenario. This implies that individuals in the sample can expect to receive more in pension benefits than what they have paid in contributions. Averaging across all birth cohorts and quintiles, under both scenarios male and female retirees can expect to receive about 1.035 and 1.054 for each euro of pension contributions, respectively. This reflects the well-known generosity of defined-benefit pension rules, which do not include any actuarial adjustment for expected post-retirement survival. Such system is particularly advantageous to individuals with low contributory levels and long retirement span. Second, IRRs decrease on average over cohorts, for both men and women. This decline in the profitability of pension contributions is consistent with longer working lives, higher contribution rates, stronger weight of the defined-contribution formula in the computation of pension annuities, and less generous definition of 'pensionable earnings' used for the calculation of pension annuities under defined-benefit criteria.[15] However, such decline in contributions' profitability is driven mostly by individuals in the bottom income quintiles, particularly in the case of men. The erosion in the progressivity of IRRs is remarkable. Under homogeneous longevity, the IRR of the bottom quintile of male cohorts 1930-1934 is about 7%, while it shrinks to about 2.7% for male cohorts 1945-1950. On the contrary, the IRR of the top quintile remains stable around 3% across all male cohorts. Similar findings apply to women. Under homogeneous longevity, the IRR of the bottom quintile of female cohorts 1930-1934 exceeds 14%, while it collapses to about 5% for cohorts 1945-1950. A modest decline in the IRR is observed also for women in the top quintile, from about 4.4% for cohorts 1930-1934 to 3.6% for cohorts 1945-1950. Third, the distributional implications of het-

[15]Starting with the 1992 reform the definition of pensionable earnings was modified so as to gradually take into account the entire working life. Before the reform, pension benefits were calculated using defined-benefit criteria with pensionable earnings defined on the basis of the individual's last years' average earnings, which implied also very high substitution rates.

erogeneous longevity for the profitability of pension contributions are sizeable in the case of men (Figure 3.15). For the bottom quintile, the IRR is significantly lower under the assumption of heterogeneous mortality as compared to the homogenous longevity scenarios, while it increases for top quintiles. For male cohorts 1930-1934, average IRR is 0.14 percentage points lower at the bottom and 0.2 percentage points higher at the top. Such differences become pronounced across cohorts. Indeed, for male cohorts 1945-1950 IRRs under heterogeneous mortality are 0.3 percentage points lower at the bottom, and 0.25 percentage points higher at the top. In relative terms, these differences correspond to about a 14% decline and 10% increase, respectively. In other words, widening disparities in life expectancy among males magnify the erosion of progressivity in the Italian pension system implied by reforms aimed at tightening the link between contributions and pension benefits. As for women, consistently with patterns documented for pension wealth, differences in IRR under homogenous and heterogeneous longevity are quantitatively negligible (± 0.05 percentage points) (Figure 3.16). In this case too, we get analogous results when investigating the evolution of IRRs by retirement year rather than by birth cohort. In interpreting these findings, it is worth recalling that our analyses do not include types of pension benefits, such as survivors, disability and social pensions, which may alter the distribution of IRRs.

Finally, to better appreciate the determinants of contributions' profitability at the individual level, we regress IRRs calculated under heterogeneous longevity against retirees' characteristics available in the INPS archives (Table 3.8). Regression analysis delivers results consistent with patterns documented above. *Ceteris paribus*, IRR decreases with quintiles of lifetime income, albeit non-linearly, with retirement age and with years of contributions. It also significantly decreases with year of birth, starting from cohorts born in the mid 1930s, and for individuals who have their pension benefits partially computed under defined-contribution pension rules (mixed regime). Being a woman or a recipient of early pension benefits is also associated with higher IRR.

3.5 Discussion

In this study, we leverage a compendium of administrative data provided by the Italian Social Security Institute to advance our knowledge about lifespan inequalities in the Italian adult and elderly population, and to evaluate their policy implications for the pension system. Our analysis delivers two main sets of findings.

First, we shed light on some worrisome trends in the evolution of lifespan inequalities among the adult and elderly population in Italy. While mortality delay (increasing average age at death) and mortality compression (declining lifespan variability) are observed across all socio-economic strata, our findings suggest that these improvements have not been equally shared, particularly in the case of men. Indeed, cohort-based analysis reveals that disparities in life expectancy among men have been widening across cohorts, in particular when measuring socio-economic status by mid-career employment income. In addition, men who belong to the upper tail of the mid-career employment income distribution and who occupy managerial positions face increasingly less uncertainty in age-at-death compared to men of opposite socio-economic status. Period-based analysis delivers qualitatively consistent results, suggesting that male retirees at the top of the pension income distribution face increasingly higher life expectancy and increasingly lower lifespan variation compared to male retirees at the bottom of the pension income distribution. Period-based findings suggest that such disparities are particularly pronounced in Northern regions, although increasing disparities in life expectancy between the top and the bottom of the pension income distribution are observed across most Italian regions. In the case of women, instead, we find no clearly discernible gradient in life expectancy and lifespan variation, and no clear trends either from a cohort- or period-perspective, particularly when measuring socio-economic status based on own income. Still, it is possible that for the cohorts and periods considered, women's own income represents a poor proxy of women's true socio-economic status. Sensitivity analyses relating female retirees' mortality to husbands' pension income suggest that spouses' income should be jointly considered for a proper assessment of lifespan inequalities among women in Italy. For this reason, occupation should be considered as a better indicator of women's socio-economic status as compared to (individual) labour and pension income, at least for the considered periods and cohorts.

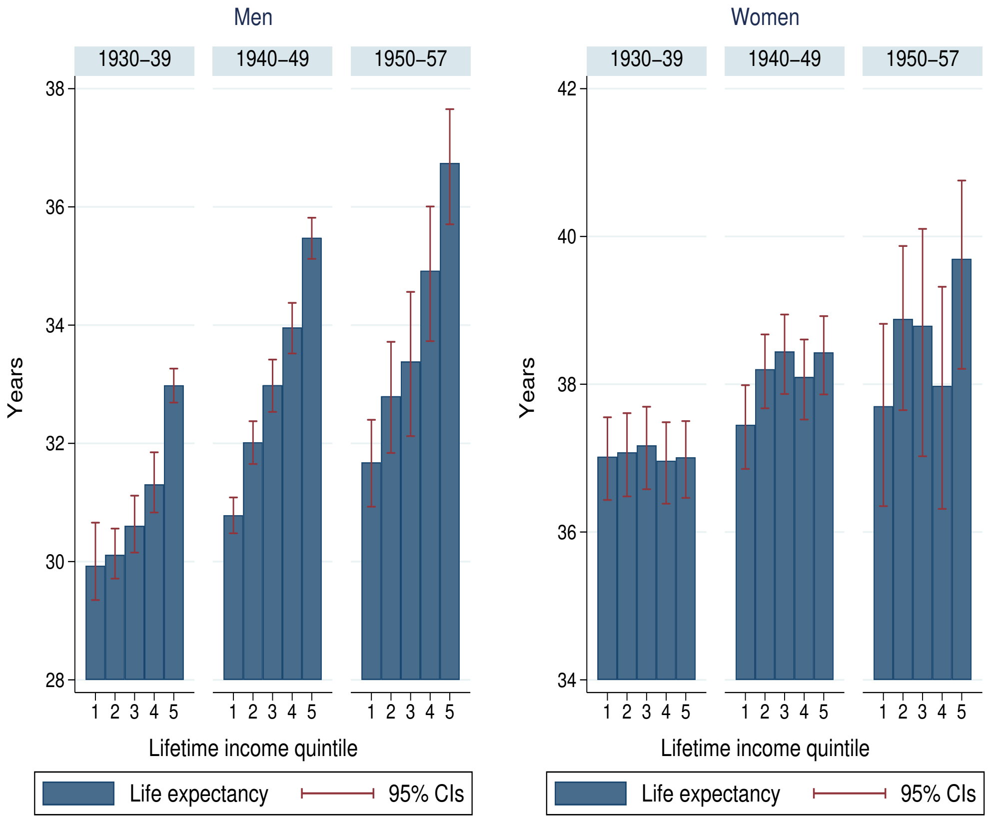
A number of caveats apply. First, comparison of period- and cohort-based mortality estimates calls for caution as they may reflect very different dynamics. Indeed, period-based analysis incorporates the effects of factors or events that *equally* affect all age groups at a particular calendar time. Cohort-based analysis, instead, incorporates the effects of factors or events that affect age groups *differetially*. Moving to data-related issues, a major limitation of our analysis is that it is based on individuals (formerly) employed in the private sector only. As such, it does not cover individuals with either a public employment or self-employment background, and individuals with no formal employment background *tout court*. This limits the generalizability of our findings to the entire Italian population. Moreover, the definition of occupational class used in this chapter, which follows the taxonomy set by Italian labour law on private employment and which has been historically used in INPS archives, is rather coarse, as each class encompasses a range of extremely heterogeneous professions. As shown in Chapter 2 of this thesis, there are tangible differences in mortality and longevity even among occupational categories belonging to the same occupational macro-class. As such, this study may fail to detect important trends in lifespan inequalities emerging around specific occupations. Third, this study does not focus on the mechanisms which may explain the observed evolution in lifespan inequalities. The lack of cause-specific mortality data prevents us from identifying explanations linked to health-related behaviours

Coming to the second set of findings, we show that distributional implications of heterogeneous longevity for the pension system are tangible, notably in the case of men. In particular, we document that the erosion in the profitability of pension contributions implied by heterogeneous longevity is stronger for male retirees at the bottom of the lifetime income distribution, and that such dynamics have become more pronounced over time. In the case of women, the lack of a clear socio-economic gradient in mortality implies that the distributional consequences are limited, too. Interpretation of these results requires some caution, though. As stressed, our analysis does not account for specific types of pension benefits, such as disability or survivor benefits, which may alter distributional dynamics. Moreover, our estimates are based on individuals who retired mostly under defined-benefit pension rules. If pension benefits were entirely computed under notional defined-contribution rules, which set a tighter link between pension benefits and contributions assuming homogeneous residual lifespan at retirement, the distributional implications of heterogeneous longevity would be even more important (Mazzaferro et al. 2012).

Tracking the evolution of lifespan inequalities, in their various dimensions, is crucial for setting equitable public policies, notably those related to social security systems. Beyond being troubling *per se*, lifespan inequalities raise a number of critical issues when it comes to policies which overlook such disparities, such as those linking statutory retirement age to changes in population average life expectancy (Alvarez et al. 2020) or linking pension benefits calculation to coefficients which are not differentiated by relevant socio-economic variables (Caselli & Lipsi 2018). This kind of measures, which have been adopted in Italy since the mid 1990s, penalize individuals with systematically less favourable survival profiles compared to that of the general population. While population ageing makes reforms for ensuring the sustainability and the *inter*-generational equity of social security systems unavoidable, widening inequalities hiding behind population ageing makes the issue of *intra*-generational fairness of such reforms equally pressing.

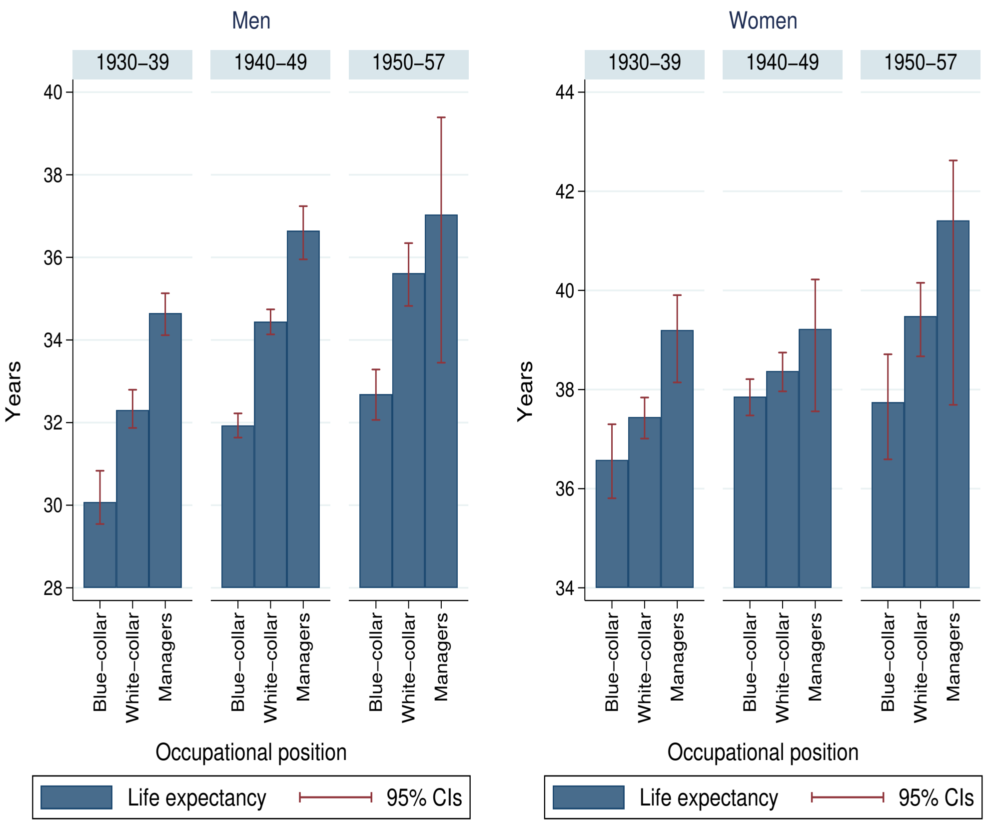
Figures Chapter 3

Figure 3.1: Life expectancy at 50 by cohort and quintiles of mid-career employment income



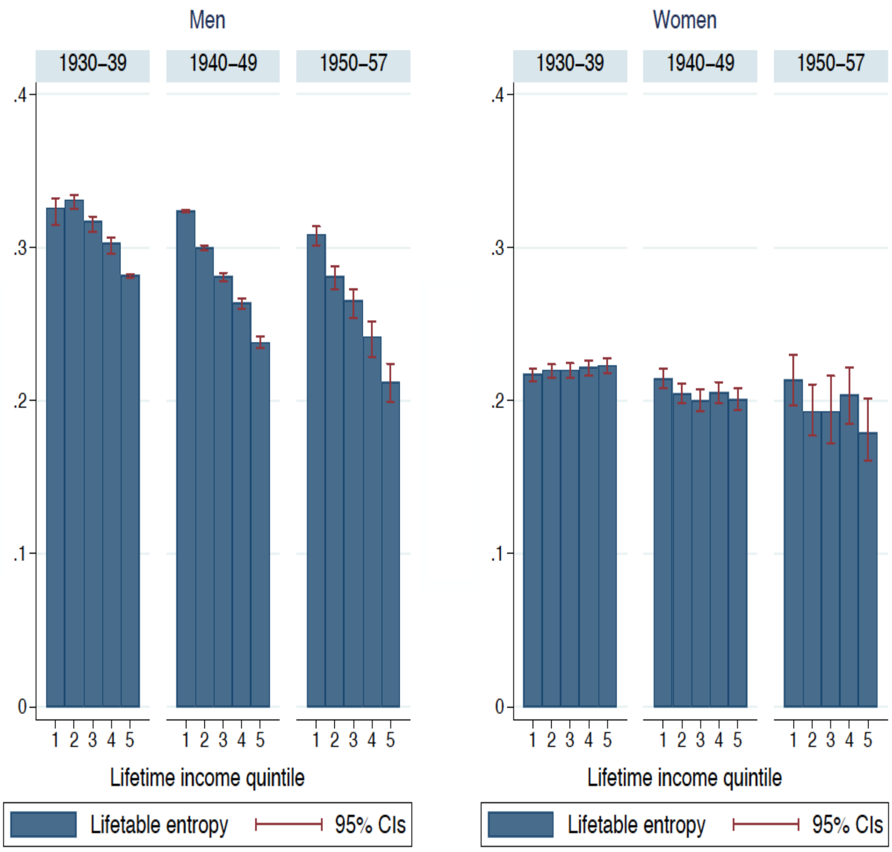
Notes: The graph plots the evolution of life expectancy at 50 by average mid-career private employment earnings, sex and birth cohort, along with 95% confidence intervals. Own elaboration based on INPS data.

Figure 3.2: Life expectancy at 50 by cohort and broad occupational group



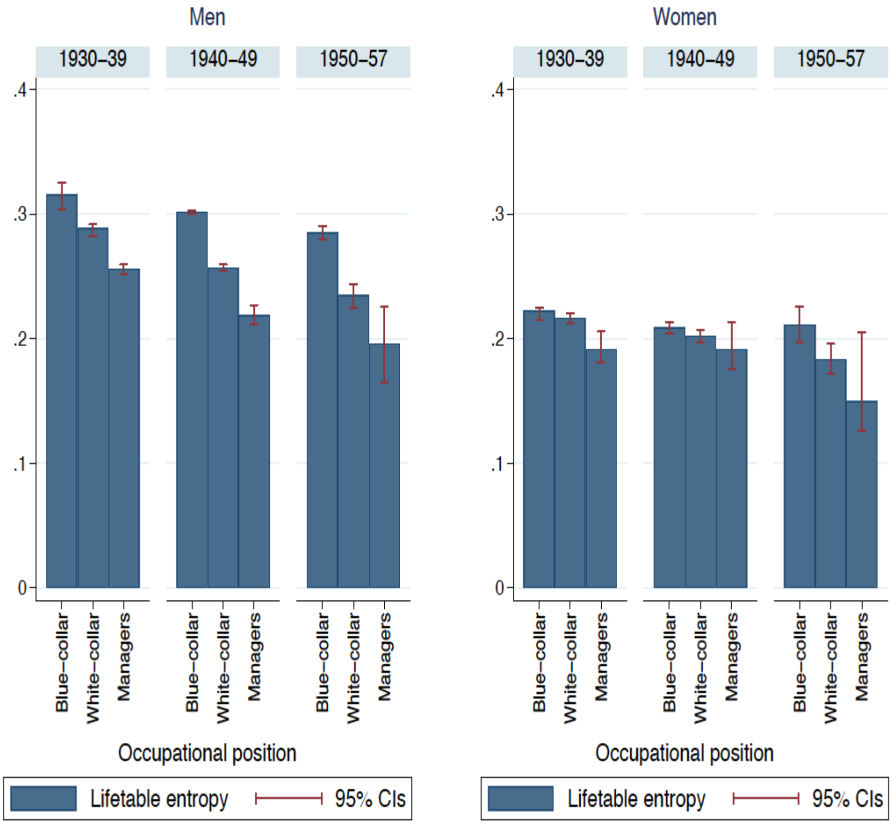
Notes: The graph plots the evolution of life expectancy at 50 by prevalent mid-career occupational position (private employees), sex and birth cohort, along with 95% confidence intervals. Own elaboration based on INPS data.

Figure 3.3: Lifetable entropy at 50 by cohort and quintiles of mid-career employment income



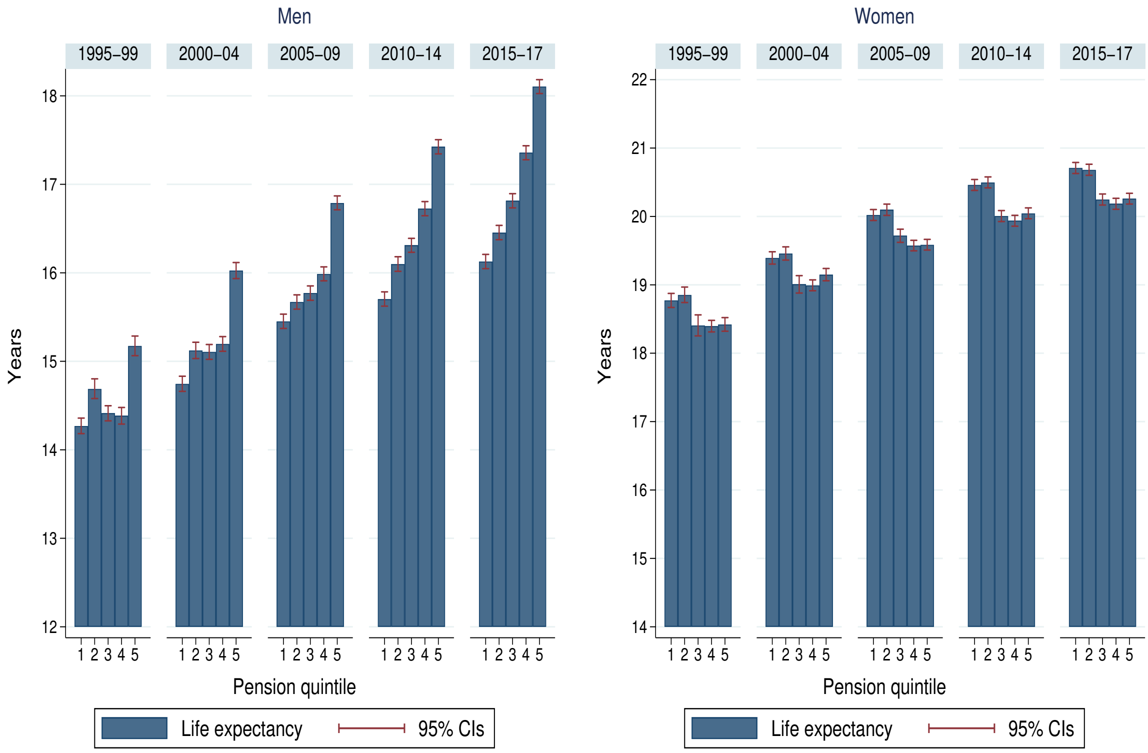
Notes: The graph plots the evolution of lifetable entropy at 50 by average mid-career private employment earnings, sex and birth cohort, along with 95% confidence intervals. Own elaboration based on INPS data.

Figure 3.4: Lifetable entropy at 50 by cohort and broad occupational group



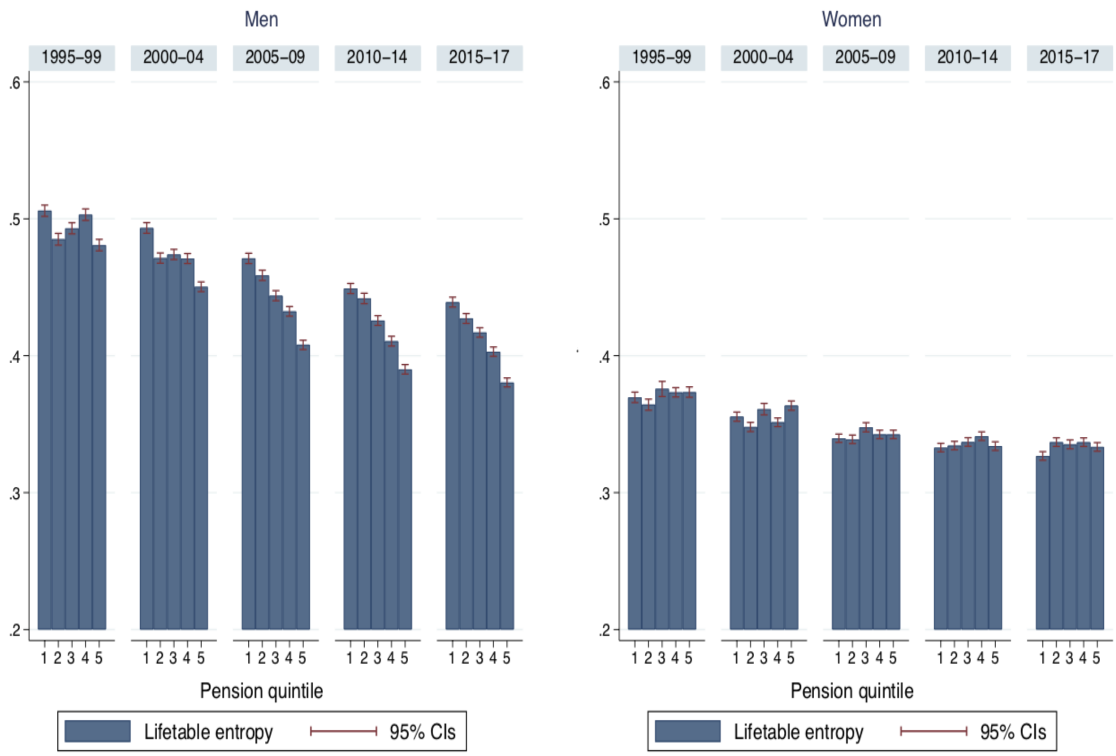
Notes: The graph plots the evolution of lifetable entropy at 50 by prevalent mid-career occupational position (private employees), sex and birth cohort, along with 95% confidence intervals. Own elaboration based on INPS data.

Figure 3.5: Life expectancy by quintiles of pension income at 67
Years 1995-2017



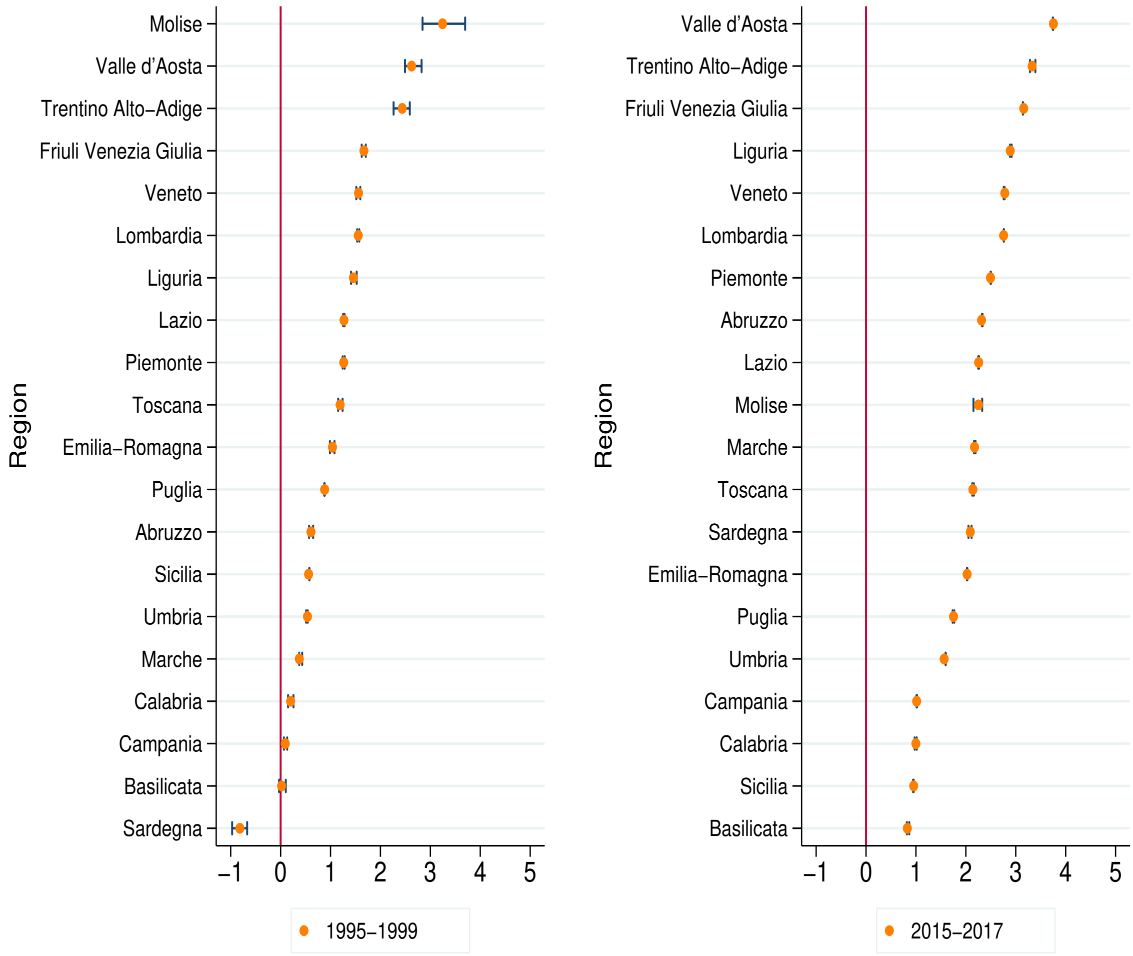
Notes: The graph plots the evolution of lifetable expectancy at 67 by pension quintile, sex and period, along with 95% confidence intervals. Own elaboration based on INPS data.

Figure 3.6: Lifetable entropy by quintiles of pension income at 67
Years 1995-2017



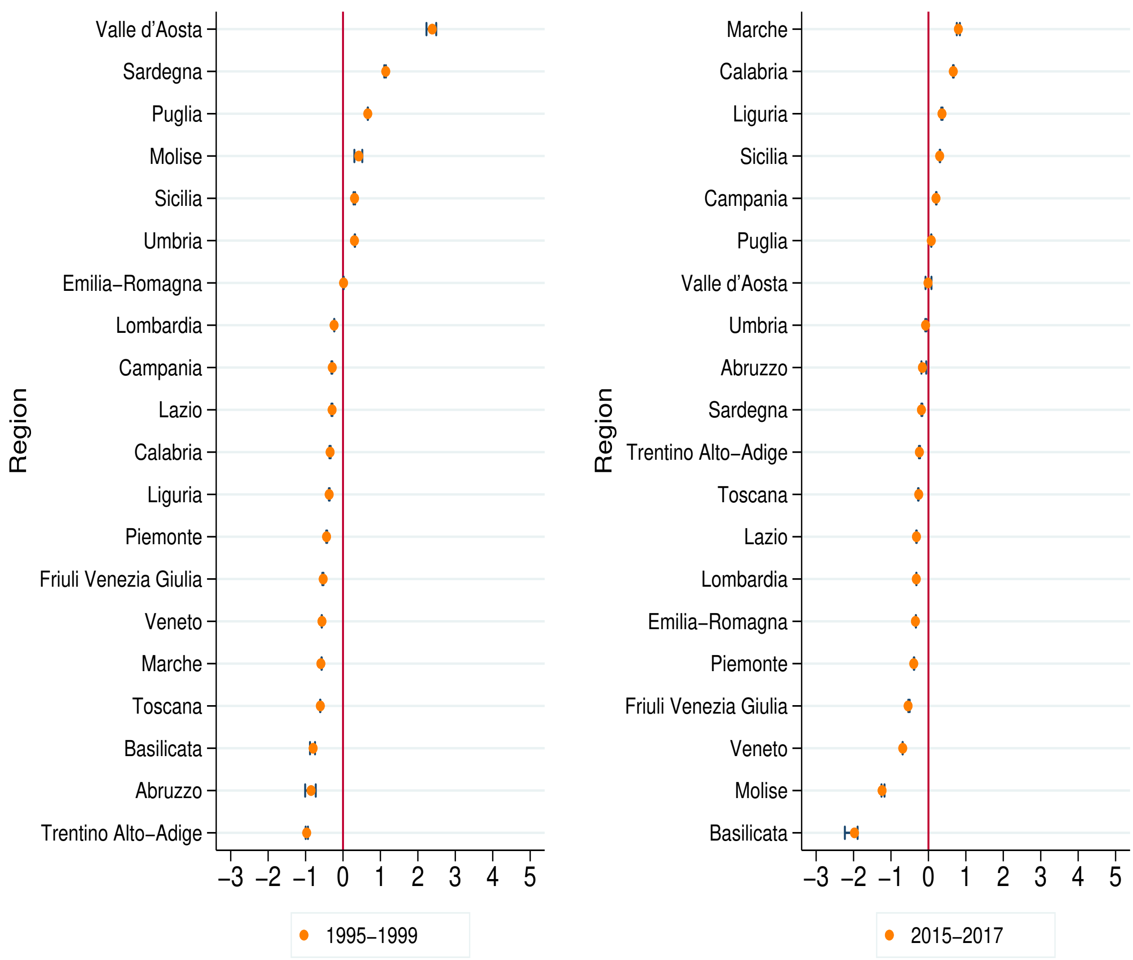
Notes: The graph plots the evolution of lifetable entropy at 67 by pension quintile, sex and period, along with 95% confidence intervals. Own elaboration based on INPS data.

Figure 3.7: Difference in life expectancy at 67 between top and bottom pension quintile
Regions - Men
Years 1995-1999 vs 2015-2017



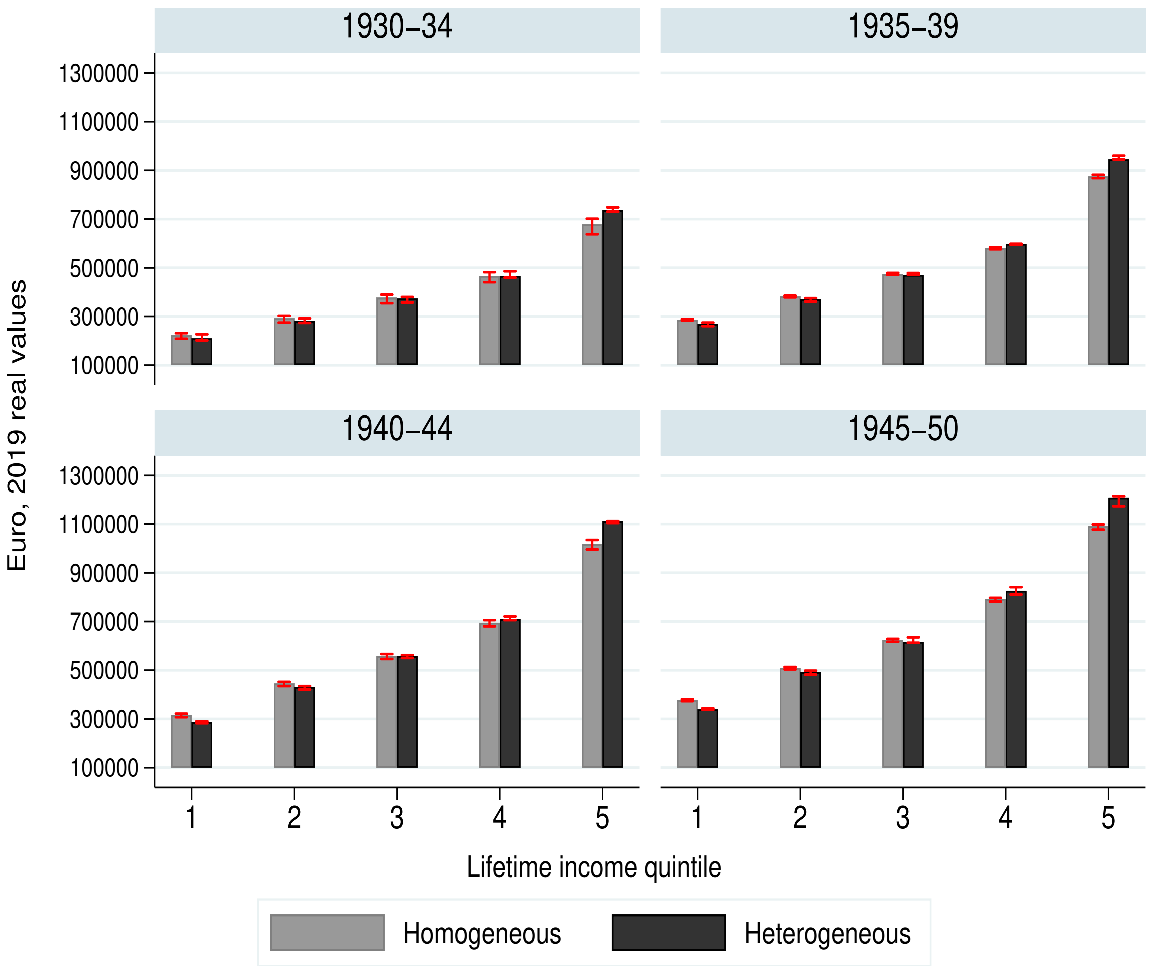
Notes: The graph plots the difference in life expectancy at 67 between top and bottom pension quintile for period 1995-1999 and 2015-2017 for each Italian region, along with 95% confidence intervals (in blue). Own elaboration based on INPS data.

Figure 3.8: Difference in life expectancy at 67 between top and bottom pension quintile
Regions - Women
Years 1995-1999 vs 2015-2017



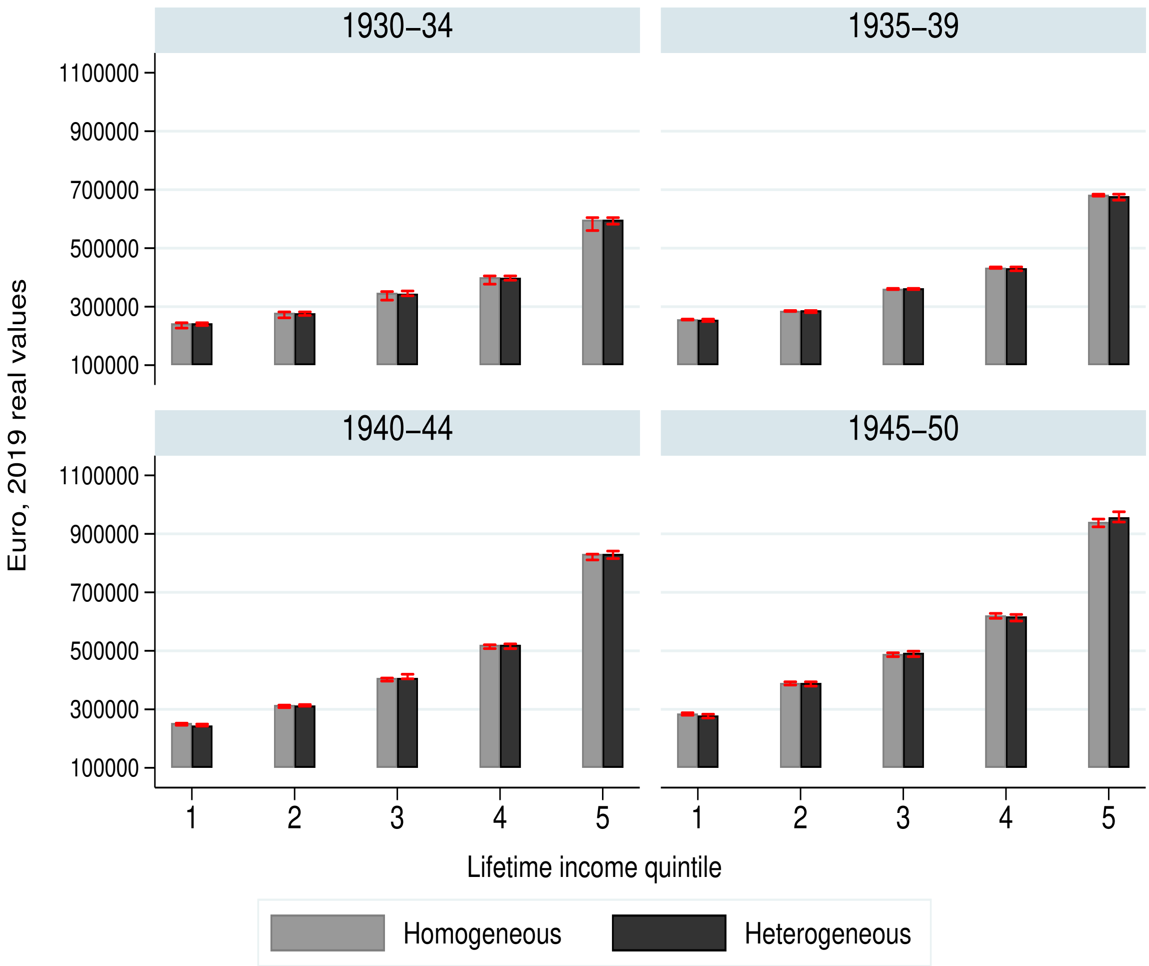
Notes: The graph plots the difference in life expectancy at 67 between top and bottom pension quintile for period 1995-1999 and 2015-2017 for each Italian region, along with 95% confidence intervals (in blue). Own elaboration based on INPS data.

Figure 3.9: Pension wealth at retirement by cohort and lifetime income quintile - Men (homogeneous vs heterogeneous mortality)



Notes: Red bars represent 95% confidence bands accounting for uncertainty in mortality estimates.

Figure 3.10: Pension wealth at retirement by cohort and lifetime income quintile - Women (homogeneous vs heterogeneous mortality)



Notes: Red bars represent 95% confidence bands accounting for uncertainty in mortality estimates.

Figure 3.11: Difference in pension wealth at retirement with and w/o heterogeneous mortality - Men

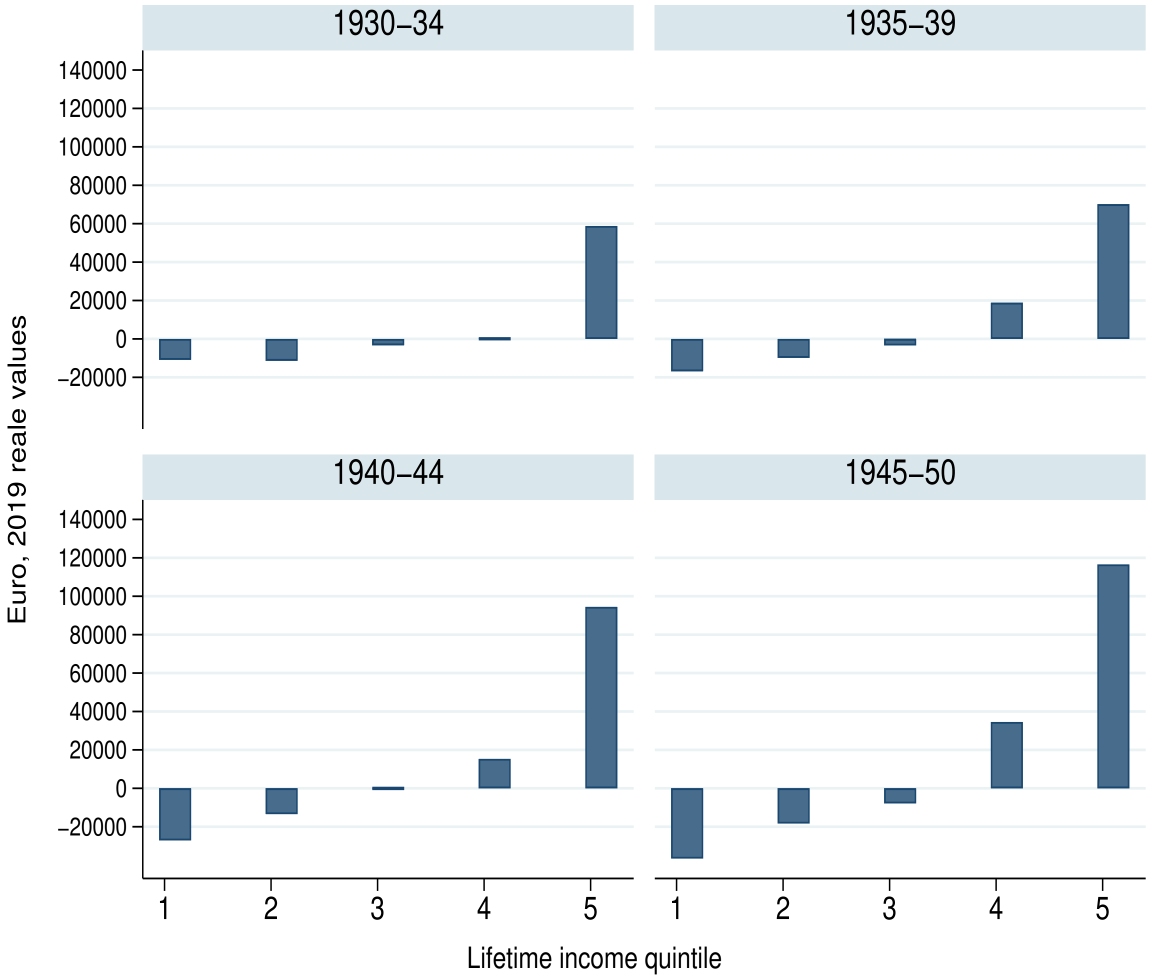


Figure 3.12: Difference in pension wealth at retirement with and w/o heterogeneous mortality - Women

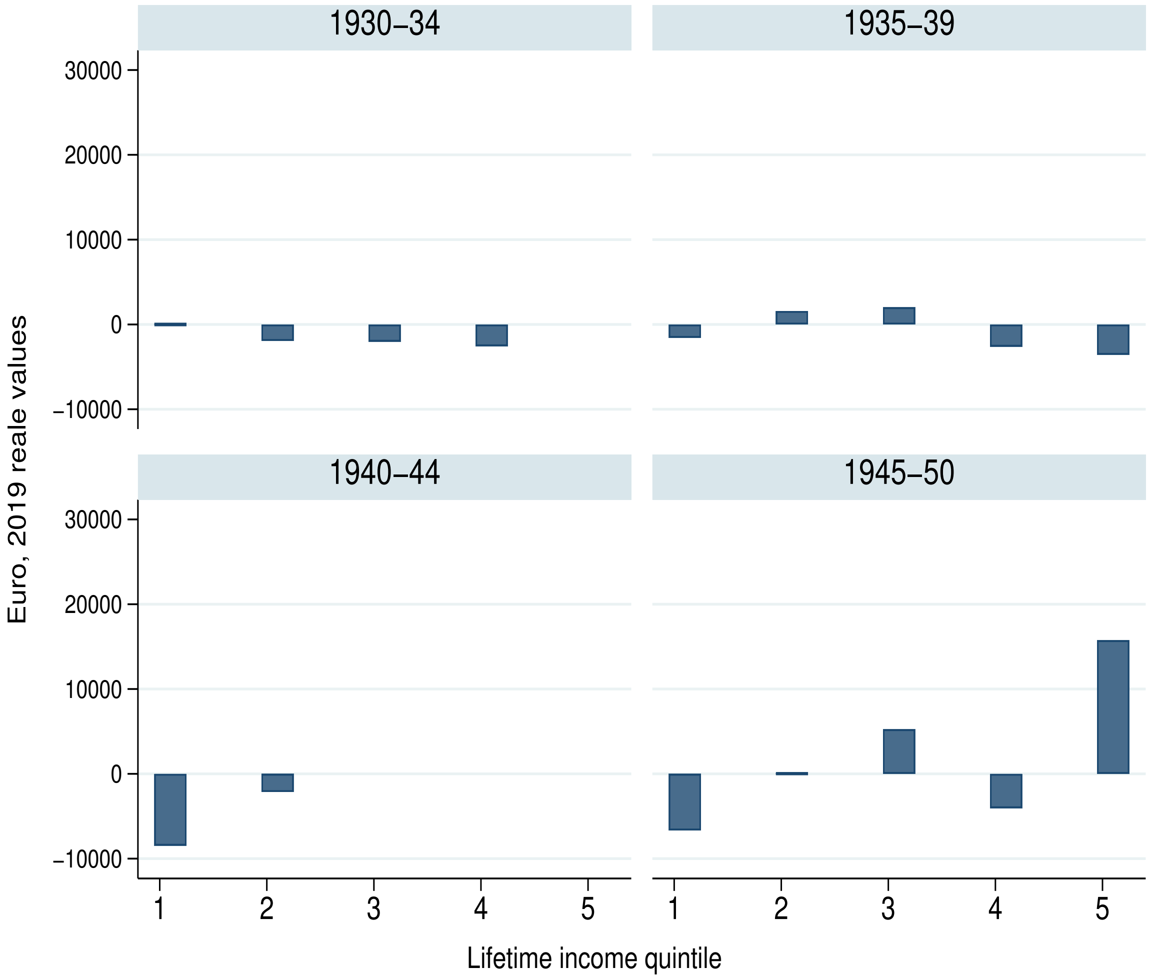
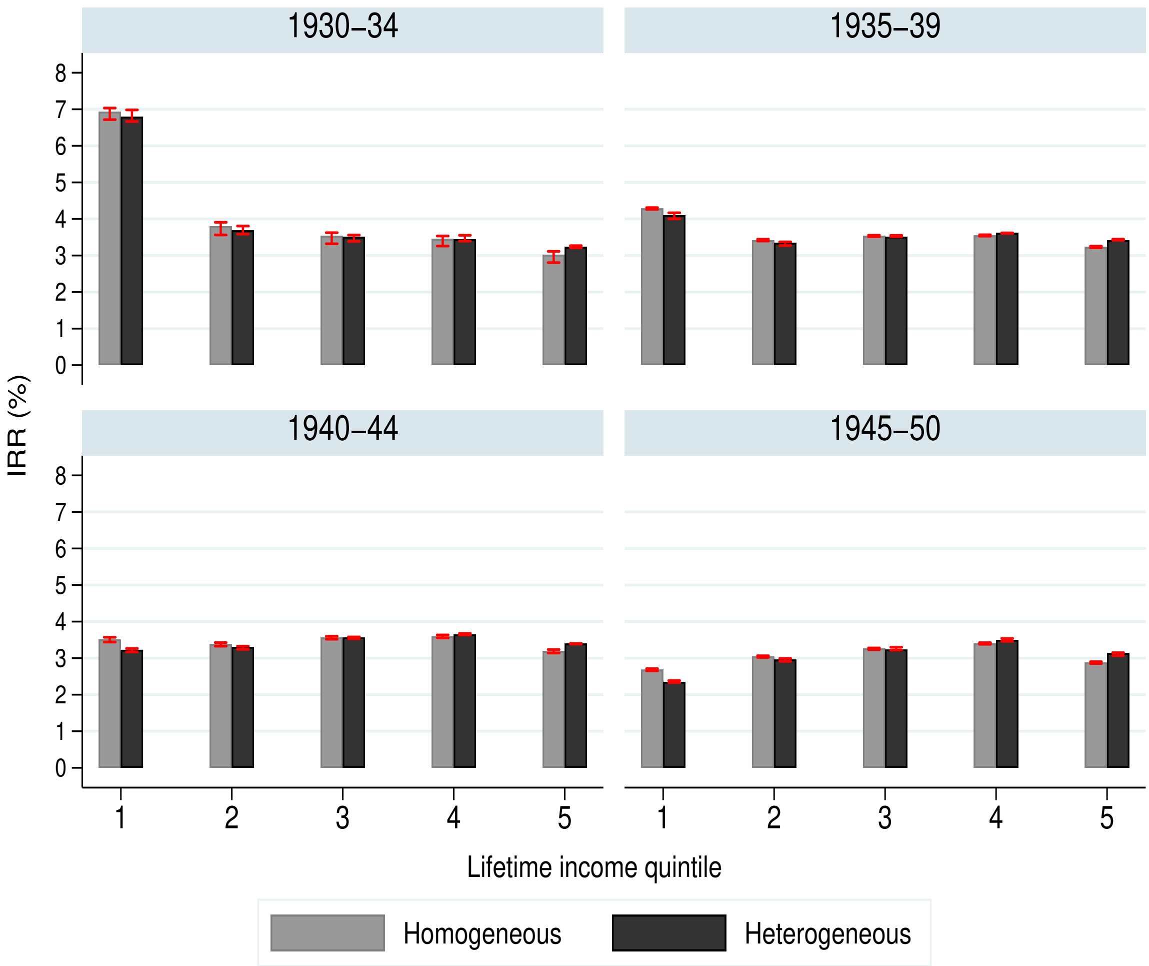
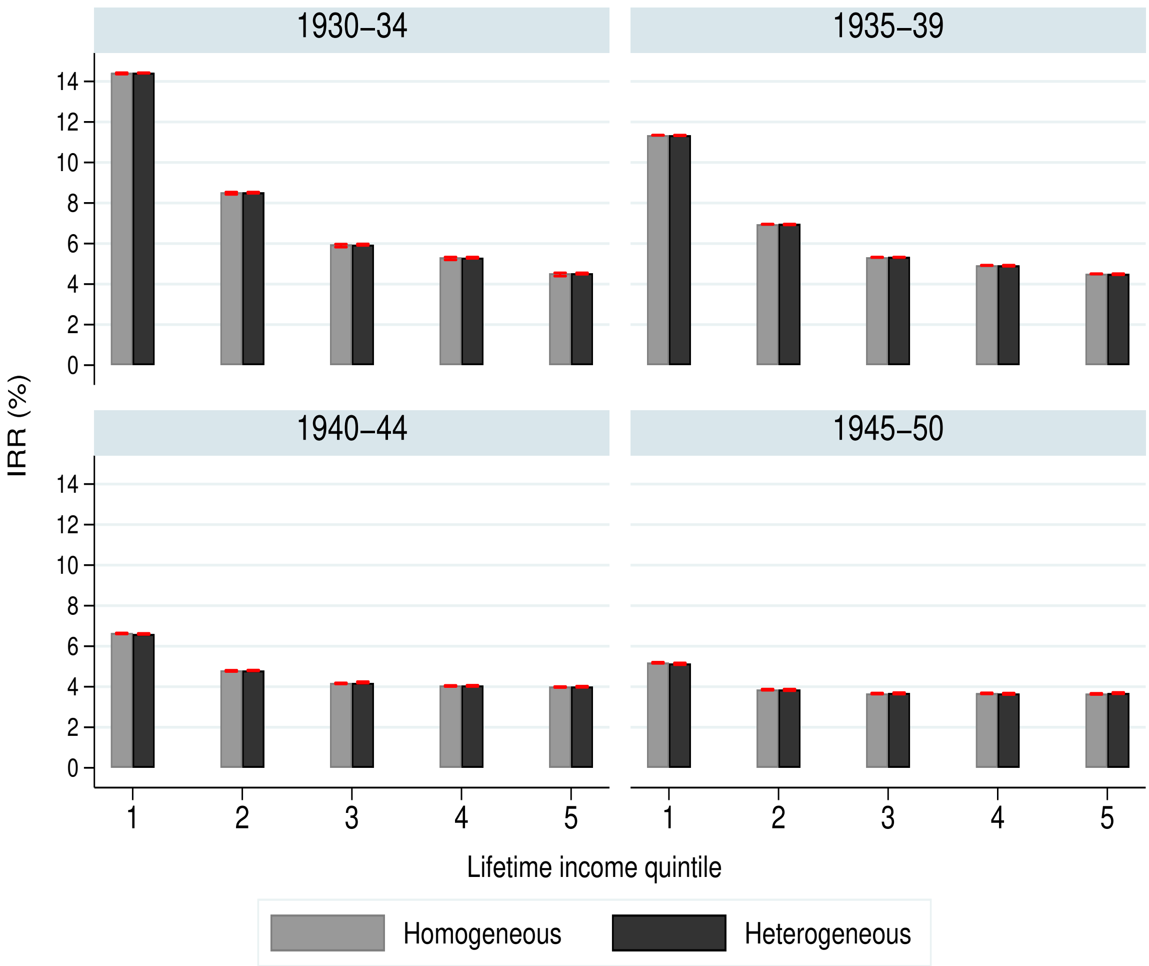


Figure 3.13: Internal rate of return by cohort and lifetime income quintile – Men (homogeneous vs heterogeneous mortality)



Notes: Red bars represent 95% confidence bands accounting for uncertainty in mortality estimates.

Figure 3.14: Internal rate of return by cohort and lifetime income quintile – Women (homogeneous vs heterogeneous mortality)



Notes: Red bars represent 95% confidence bands accounting for uncertainty in mortality estimates.

Figure 3.15: Internal rate of return by cohort and lifetime income quintile – Men (homogeneous vs heterogeneous mortality)

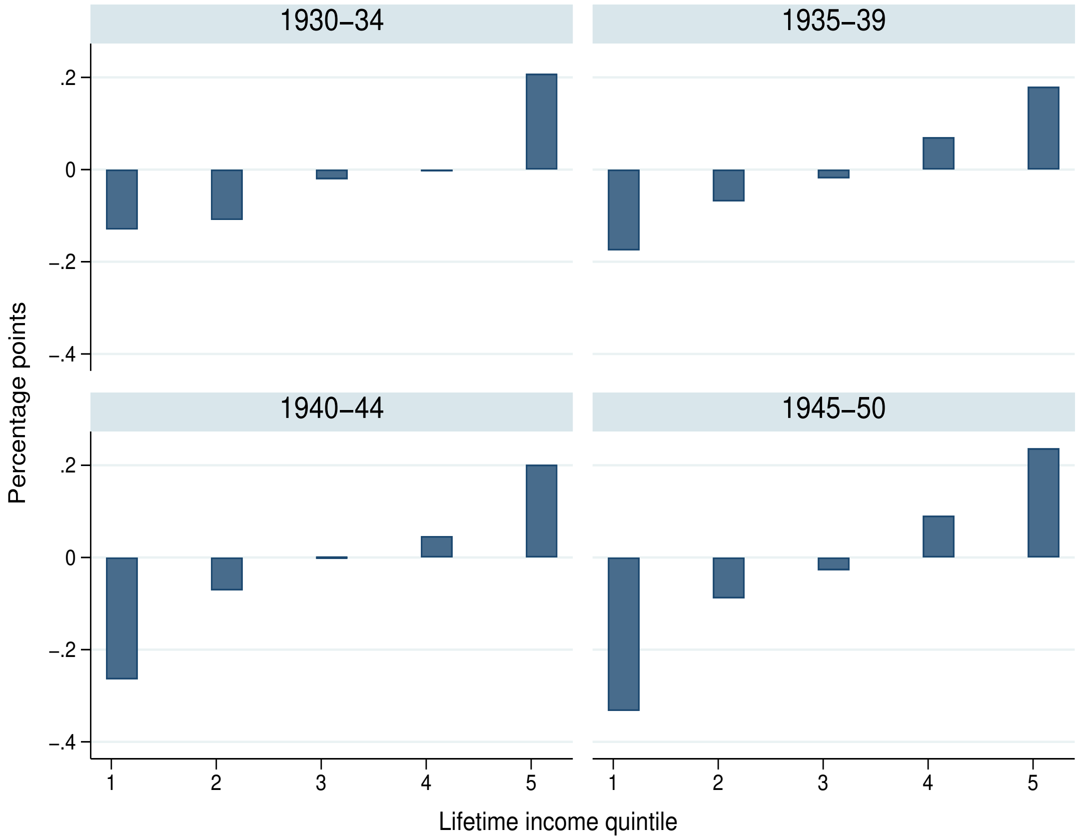
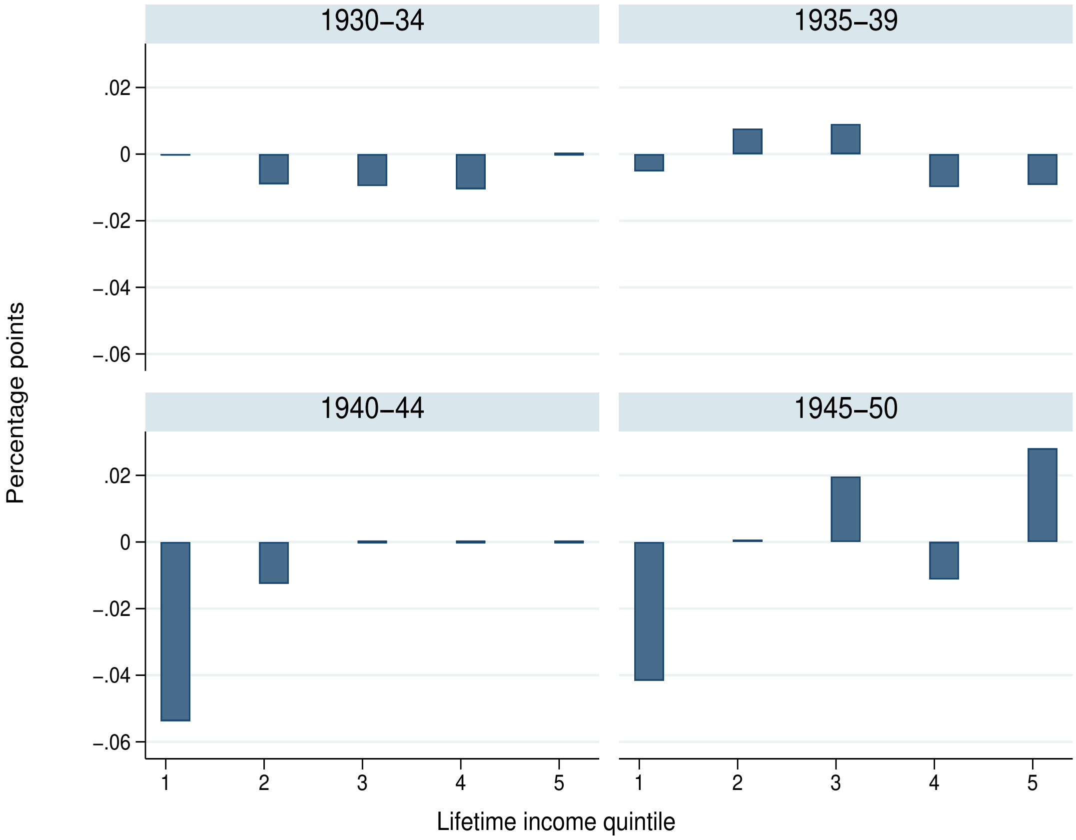


Figure 3.16: Internal rate of return by cohort and lifetime income quintile – Women (homogeneous vs heterogeneous mortality)



Tables Chapter 3

Table 3.1: Construction of the *Dichiarazioni Uniemens* sample

	N. of individuals
Starting sample	21,966,659
Dropping individuals born prior to 1930 or after 1957	14,909,172
Dropping individuals who die before 50	44,229
Dropping individuals who retire before 50	64,012
Final sample	6,949,246
Final sample men	4,842,306
Final sample women	2,106,940

Table 3.2: *Dichiarazioni Uniemens* sample
Observations by year of birth and survival as of 2018

Year of birth	Not dead	Dead	Total
1930	84,650	160,286	244,936
1931	91,959	147,374	239,333
1932	99,232	135,492	234,724
1933	110,470	129,806	240,276
1934	120,000	119,003	239,003
1935	131,024	109,624	240,648
1936	132,594	95,077	227,671
1937	145,496	88,668	234,164
1938	161,377	85,105	246,482
1939	168,150	76,769	244,919
1940	171,667	69,184	240,851
1941	160,376	57,076	217,452
1942	161,946	51,392	213,338
1943	160,420	45,724	206,144
1944	164,204	40,805	205,009
1945	159,569	35,021	194,590
1946	214,160	40,940	255,100
1947	218,456	35,896	254,352
1948	228,316	33,320	261,636
1949	230,081	29,298	259,379
1950	232,994	26,598	259,592
1951	234,950	23,095	258,045
1952	240,056	20,928	260,984
1953	249,935	19,013	268,948
1954	265,463	17,590	283,053
1955	277,681	15,993	293,674
1956	291,203	14,597	305,800
1957	305,807	13,336	319,143
Total	5,212,236	1,737,010	6,949,246

Table 3.3: *Casellario Pensionati* sample
Observations by calendar year and pension regime

	Men			Women		
Calendar year	Defined-benefit	Mixed	Total	Defined-benefit	Mixed	Total
1995	1,151,882	0	1,151,882	1,082,289	0	1,082,289
1996	1,219,103	84	1,219,187	1,147,169	224	1,147,393
1997	1,295,201	259	1,295,460	1,220,438	803	1,221,241
1998	1,361,474	653	1,362,127	1,289,098	1,773	1,290,871
1999	1,422,691	1,282	1,423,973	1,354,721	2,998	1,357,719
2000	1,476,454	2,513	1,478,967	1,416,304	4,379	1,420,683
2001	1,525,513	9,171	1,534,684	1,482,811	6,546	1,489,357
2002	1,554,662	22,235	1,576,897	1,545,931	9,254	1,555,185
2003	1,568,684	32,977	1,601,661	1,598,854	12,968	1,611,822
2004	1,589,493	43,689	1,633,182	1,658,625	17,706	1,676,331
2005	1,612,341	55,973	1,668,314	1,715,611	25,278	1,740,889
2006	1,634,862	69,16	1,704,022	1,745,741	58,631	1,804,372
2007	1,653,253	82,944	1,736,197	1,737,001	121,318	1,858,319
2008	1,658,256	95,267	1,753,523	1,711,853	169,539	1,881,392
2009	1,658,801	107,48	1,766,281	1,684,505	215,542	1,900,047
2010	1,657,874	121,105	1,778,979	1,654,131	260,326	1,914,457
2011	1,658,079	134,69	1,792,769	1,621,687	304,348	1,926,035
2012	1,650,774	147,859	1,798,633	1,583,819	346,047	1,929,866
2013	1,668,696	167,582	1,836,278	1,557,399	401,33	1,958,729
2014	1,683,656	186,649	1,870,305	1,529,429	453,63	1,983,059
2015	1,693,572	211,949	1,905,521	1,492,706	511,126	2,003,832
2016	1,711,681	237,729	1,949,410	1,459,618	565,071	2,024,689
2017	1,702,947	264,531	1,967,478	1,418,551	617,323	2,035,874
Total	35,809,949	1,995,781	37,805,730	34,708,291	4,106,160	38,814,451

Table 3.4: Women's post-retirement mortality and husbands' pension

	(1) Odds ratio
Woman's pension quintile [ref: 1st, bottom]	.
2nd	1.001
	(0.00959)
3rd	1.050***
	(0.0103)
4th	1.049***
	(0.0107)
5th (top)	1.165***
	(0.0129)
Husband's pension quintile [ref: 1st, bottom]	
2nd	0.987
	(0.00941)
3rd	0.997
	(0.00956)
4th	0.957***
	(0.00951)
5th (top)	0.881***
	(0.00945)
Constant	6,092***
	(9,123)
Observations	7,805,622

Notes. Results from logistic survival analysis based on female retirees from the FPLD fund, who retired between 1995 and 2017, whose husband was alive in 1995 and also retired between 1995 and 2017. The mortality follow-up extends from the year women turn 67 to the end of 2018 or the year of their death, if the latter occurs earlier. Dependent variable is a dummy taking value 1 if the woman dies by the end of the year, 0 otherwise. Pension quintiles are cohort-specific for both women and husbands. Control variables: year of birth, age difference with respect to husband, widowhood status, macro-region of residence, macro-region of birth, and 23 duration dummies.

Table 3.5: Sample construction - *Estratti Conto*

	Women	Men	All	Dropped
Initial sample	109,515	151,069	260,584	
Dropping N with no measurable employment income at 45-49	30,976	73,053	104,029	-156,555
Dropping contributions to funds other than FPLD	30,975	73,036	104,011	-18
Dropping observations prior to 1974	30,968	72,961	103,929	-82
Dropping observations with zero or missing income	30,966	72,937	103,903	-26
Dropping N with just one observation after 1974	30,894	72,796	103,690	-213
Dropping N with gaps >5 years	28,180	69,141	97,321	-6,369
Sample used to estimate lifetime earnings profile	28,180	69,141	97,321	
Dropping N who die prior to 65	27,465	65,795	93,260	-4,061
Sample used for distributional analysis	27,465	65,795	93,260	

Table 3.6: Regression coefficients (lifetime earnings profile)

	(1)	(2)
	Men	Women
Age	0.199***	0.202***
	(0.006)	(0.013)
Age^2	-0.004***	-0.001***
	(0.000)	(0.000)
Age^3	0.000***	-0.000
	(0.000)	(0.000)
Years of contributions	-0.002	-0.128***
	(0.004)	(0.009)
Constant	6.234***	4.267***
	(0.110)	(0.257)
Observations	1,484,456	55,2550
N individuals	69,141	28,180
R-squared	0.207	0.191

Results from OLS regressions with individual fixed effects. Standard errors in parentheses.
*** p-value <0.001. ** p-value <0.01. * p-value <0.05.

Table 3.7: Gini of pension wealth at retirement (Euro, 2019 real values)

Panel A: Men			
Cohort	Homogeneous mortality	Heterogeneous mortality	Difference
1930-1934	0.2736	0.2927	0.0192
1935-1939	0.2709	0.2897	0.0188
1940-1944	0.2550	0.2784	0.0235
1945-1950	0.2262	0.2554	0.0292
Panel B: Women			
Cohort	Homogenous mortality	Heterogeneous mortality	Difference
1930-1934	0.2514	0.2521	0.0007
1935-1939	0.2748	0.2737	-0.0011
1940-1944	0.2886	0.2910	0.0024
1945-1950	0.2613	0.2651	0.0038

Table 3.8: Individual determinants of internal rate of return (IRR)

	(1)
Woman	0.893***
	(0.016)
Age at retirement	-0.189***
	(0.002)
Years of contribution	-0.202***
	(0.002)
Anticipated pension	0.350***
	(0.016)
Mixed retirement regime	-1.713***
	(0.042)
Quintile of lifetime earnings [Ref: 1st (bottom)]	
Quintile 2	-1.642***
	(0.036)
Quintile 3	-1.884***
	(0.035)
Quintile 4	-1.800***
	(0.034)
Quintile 5 (top)	-1.702***
	(0.033)
Macro-region of residence [Ref: Centre]	
Abroad	-1.484***
	(0.243)
North-East	0.053*
	(0.021)
North-West	-0.013
	(0.018)
South	0.002
	(0.025)
Birth cohort [Ref: 1930]	
Birth cohort 1931	0.169**
	(0.057)
Birth cohort 1932	0.154**
	(0.052)
Birth cohort 1933	0.127*
	(0.052)
Birth cohort 1934	0.161**
	(0.053)
Birth cohort 1935	0.090
	(0.052)
Birth cohort 1936	-0.052
	(0.050)
Birth cohort 1937	-0.092
	(0.050)
Birth cohort 1938	-0.171***
	(0.048)
Birth cohort 1939	-0.232***
	(0.047)
Birth cohort 1940	-0.242***
	(0.044)
Birth cohort 1941	-0.290***
	(0.044)

Continued on next page

Table 3.8 Continued from previous page

	(1)
Birth cohort 1942	-0.314***
	(0.043)
Birth cohort 1943	-0.395***
	(0.044)
Birth cohort 1944	-0.432***
	(0.043)
Birth cohort 1945	-0.377***
	(0.045)
Birth cohort 1946	-0.412***
	(0.044)
Birth cohort 1947	-0.420***
	(0.043)
Birth cohort 1948	-0.410***
	(0.045)
Birth cohort 1949	-0.347***
	(0.049)
Birth cohort 1950	-0.305***
	(0.050)
Constant	22.843***
	(0.163)
Observations	93,260
R-squared	0.429

Notes. Notes. Results from OLS regressions. Robust standard errors in parentheses. *** p-value <0.001. ** p-value <0.01. * p-value < 0.05.

Chapter 4

Putting COVID-19 into historical perspective: was it just a flu?

4.1 Introduction

Since its emergence in Asia in late 2019, the coronavirus COVID-19 pandemic has ravaged societies around the world. However epochal, though, COVID-19 is hardly the first pandemic human societies were confronted with during the last few decades (Snowden 2019). In the second half of the 20th century, two prominent flu pandemics originating in Asia (the 1957-58 and 1968-1970 pandemics) hit the world, causing important human losses worldwide (Viboud et al. 2005; 2013). Despite their demographic relevance, the mortality impact of these major health shocks has not been investigated thoroughly yet, particularly at the subnational level. Subnational focus is central to evaluate the mortality burden of flu pandemics. Indeed, if transmission is geographically clustered in nature, as in the case of COVID-19, country-level analyses are likely to yield a dampened picture of the mortality impact of major influenza episodes (Blangiardo et al. 2020; Trias-Llimós et al. 2020). Still, studies which seek to quantify and compare the mortality impact of major past influenza outbreaks by accounting for the spatial nature of pandemic waves are rare (Chowell et al. 2014). The study presented in this chapter contributes to address

[1] This chapter is the result of a joint work with S. Ghislandi. The analysis on the mortality impact of the COVID-19 pandemic in Italian provinces presented in this chapter subsumes and updates part of the results of a joint work with S. Ghislandi, R. Muttarak and M. Sauerberg, accepted for publication in the *Vienna Yearbook of Population Research* under the title "Human costs of the first wave of the COVID-19 pandemic in the major epicenters in Italy".

this gap by investigating the impact of post-World War II major flu pandemics on mortality patterns in Italy at the provincial level based on a newly constructed digital database with age-sex-province specific mortality and population statistics spanning nearly three decades (1953-2020). Specifically, we estimate and compare the human costs associated to the 1957-1958 and the 1968-1970 pandemics in Italian provinces with the provisional death toll of COVID-19 in the same areas, which allows us to put the latter into historical perspective.[1] To the best of our knowledge, ours is the first attempt to compare the mortality burden of COVID-19 at the subnational level, in a hardly hit country, i.e. Italy, with that of past influenza pandemics.

To quantify the mortality burden of the 1957-1958 and 1968-1970 flu pandemics, and to make it comparable with that of the COVID-19 pandemic, we adopt an all-cause excess mortality approach. All-cause excess mortality compares mortality levels registered in a given place and time, for all types of causes, to what would have been normally expected. This approach has been previously used to estimate the global mortality burden of the Spanish flu (Murray et al. 2006). Recently, it has been extensively employed to assess the mortality impact of COVID-19 in a number of countries, both at the national and sub-national level (Aburto, Schöley, et al. 2021; Aburto et al. 2020; Blangiardo et al. 2020; Modig et al. 2021; Trias-Llimós et al. 2020). Focusing on all-cause excess mortality allows to overcome a number of empirical issues. First, data on influenza-related fatalities may not be available or be imprecisely collected, both at the national and sub-national level. Historical and geographical comparison can be particularly problematic as public authorities may have different collection, recording and reporting policies regarding influenza cases and casualties. Second, influenza-specific mortality data, when available, fail to capture the effect of influenza outbreaks on other causes of deaths due, for instance, to depletion of hospital capacity or disruptions in routine medical treatments.

[1]Provinces in Italy represent the intermediate administrative tier between municipalities and regions.

While excess mortality is undoubtedly useful to measure the mortality burden of pandemic events, it needs to be translated into metrics which allow for reliable and consistent comparison of mortality shocks over time and place. In this regard, life expectancy at birth, which measures the average longevity of a hypothetical group of newborns that experience the death rates observed in a given period over their entire life course, is a suitable candidate. While in reality individuals cannot be expected to experience such death rates over their life course, life expectancy has a number of useful properties. First, it is significantly related to the overall well-being of a population (Ghislandi et al. 2019; Sen 1988). Second, it is not influenced by the age structure of populations under study. As such, it does not depend on the choice of any arbitrary population standard as in the case of standardized deaths rates or years of life lost (Marois et al. 2020). Moreover, being it sensitive to ages at which deaths occur, it enables effective comparison of mortality shocks affecting populations along different age groups.

Lifespan variation is an alternative metric to quantify and qualify dynamics in population health and mortality (van Raalte et al. 2018). Albeit related, life expectancy and lifespan variation capture two different dimensions of mortality (Aburto et al. 2020). While life expectancy summarizes the average age at death of individuals belonging to a given population, lifespan variation expresses the variability in age-at-death between individuals within that population. Two populations may thus share the same life expectancy while displaying different levels of lifespan variation due to differences in the age distribution of deaths. The mortality dynamics which underpin improvements in both indicators are also different. Reductions in mortality rates at any age lead always to increases in life expectancy. For lifespan variation to decrease when life expectancy is increasing, though, more deaths need to be averted at younger ages as compared to older ages, the threshold age corresponding generally to life expectancy. In other words, mortality compression, i.e. deaths occurring at an increasingly similar age, should be observed. For these reasons, lifespan variation is a powerful tool to explore how (un)evenly health and mortality shocks are distributed within a given population. It is worth stressing that while life expectancy tends to decrease in times of pandemic, due to increased mortality rates across the entire age spectrum (Viboud et al. 2005), lifespan variation may increase or decrease depending on the age distribution of excess deaths (Aburto, Schöley, et al. 2021). Given their useful properties, both life expectancy and lifespan variation have been widely used to quantify the mortality burden of COVID-19, both at the national (Aburto, Kashyap, et al. 2021; Aburto, Schöley, et al. 2021) and subnational level (Trias-Llimós et al. 2020). A further joint advantage of both measures is that mortality data, from which both metrics are derived, are available for some countries since the early nineteenth century. This makes it possible to use life expectancy and lifespan variation for historical comparisons of major mortality shocks.

Building on these considerations, in this study we investigate the mortality burden of the 1957-1958 and 1968-70 pandemics in Italian provinces by looking at all-cause excess mortality, and at the impact on both life expectancy and lifespan variation, measuring the latter in terms of standard deviation in age-at-death. We then contrast such mortality patterns, and the related human costs, with those documented for the COVID-19 in the first year of the pandemic (i.e. calendar year 2020) over the same geographical areas. The remainder of the chapter is organized as follows. In Section 4.2, we provide a brief historical overview of the three pandemics under study. In Section 4.3, we illustrate the data and the methods we employ. In Section 4.4, we present the results, and Section 4.5 concludes.

4.2 Background

Historical studies on prominent influenza outbreaks of the 20th century have concentrated overwhelmingly on the Spanish flu. Having killed approximately 50 million people worldwide between 1918 and 1921 (Johnson & Mueller 2002; Murray et al. 2006), the Spanish flu was as a major health shock with long-lasting consequences on human societies, from population health (Noymer & Garenne 2000) to human capital development (Almond 2006) and trust (Aassve et al. 2021). Relatively little attention has been devoted, instead, to flu pandemics which followed in the ensuing decades and for which humans possessed limited or no immunity: the 1957-1958 pandemic (also known as the 'Asian flu') and the 1960-1970 pandemic (also known as the 'Hong Kong flu').

The influenza strain (H2N2) responsible for the the 1957-1958 pandemic was first detected in the Yunnan province in China, in February 1957 (Saunders-Hastings & Krewski 2016). The virus spread through Hong Kong, Singapore, Taiwan and Japan, before spreading globally in the summer of 1957. The contagion rates soared markedly in the early fall of 1957, with the re-opening of schools (Henderson, 2009). Despite the limited diffusion of vaccines and the lack of vigorous non-pharmaceutical interventions, the global mortality burden was comparatively contained, with one to two million estimated deaths worldwide (Saunders-Hastings & Krewski 2016). A study by Viboud et al. (2013), based on data from 39 countries encompassing Europe, the Asian-Pacific region and the Americas, estimates that total excess deaths in 1957-1959 totaled to 1.1 million, with an average excess respiratory mortality rate of 1.9 per 10,000 people. Age-specific excess mortality rates were highest among infants (0-4 years old) and elderly (>65 years old). In Italy, a report released by the High Commissariat for Hygiene and Public Health in May 1958 estimated that the Asian flu killed around 30,500 people between August 1957 and February 1958.[2] As documented by local newspapers, public health responses included the postponement of school opening and the temporary closure of schools in the fall of 1957, albeit limited to some areas of the country.[3]

[2] Source: *Il Corriere della Sera*, 8 May 1958, p. 9

[3] Source: *L'Unità*, 4 October 1957, p. 2

About ten years after its first appearance, the Asian flu strain underwent an antigenic drift emerging as a new virus (A/H3N2) which was officially reported for the first time in Hong Kong in July 1968 (Henderson et al. 1969). The following month, the virus was isolated in Japan and in the United States, the spread being possibly driven by veterans repatriating from the Vietnam War. In September, cases were reported in Great Britain. In early 1969, the virus was isolated in continental Europe, with infections being detected in France. The pandemic is estimated to have called between 500,000 and 2 million lives worldwide between 1968 and 1970 (Henderson, 2009). As reported by Viboud et al. (2005), the pandemic hit through two distinct waves. The first one (1968-69) took its toll mostly in the United States, while the second one (1969-1970) killed mostly in Europe. Non-pharmaceuticals interventions, including school closure and quarantine measures, were limited, public health response consisting mostly in the hospitalization of most serious cases (Henderson, 2009). In Italy, the pandemic hit in the winter of 1969-70 (Ragona et al. 1978). As reported by local newspapers, the surge in cases put hospitals under pressure[4], forcing the temporary closure of schools and the reduction of public transportation services in some areas of the country due to personnel's sickness absenteeism.[5] Overall, the Hong Kong flu is estimated to have killed between 20,000[6] and 57,000 people in Italy, with excess deaths recorded across all age groups, especially the 45-64 group (Rizzo et al. 2007).

Fifty years after the 1968-1970 pandemic, in December 2019 a cluster of cases of pneumonia of unknown aetiology was reported in Wuhan, in the Hubei Province, China. In early January 2020, the Chinese Center for Disease Control and Prevention reported that a novel coronavirus (SARS-Cov-2) was identified as causative agent for some cases of pneumonia. While the city of Wuhan was locked down, by the end of January im-

[4]Source: *Il Corriere della Sera*, 29 December 1969, p. 2

[5]Source: *Il Corriere della Sera*, 9 December 1969, p. 17

[6]Estimate reported by the Italian National Institute of Health. See https://www.epicentro.iss.it/passi/storiePandemia

ported cases of the novel coronavirus 2019 (COVID-19) from China were reported in a number of countries in Asia (Japan, Thailand, South Korea), Europe (France, Germany) and America (United States). In Europe, the first severe case of local transmission of COVID-19 was diagnosed on 21 February at a small hospital in Codogno, a municipality in the province of Lodi, south-east of Milan, Italy (Paterlini 2020). Over the following days, the Italian authorities reported clusters of cases in several regions in the North of the country (Lombardy, Veneto, Piedmont, Emilia-Romagna). A couple of weeks later, the Italian government introduced strict public health measures, imposing a partial nationwide lock-down, followed by a total lockdown of all non-essential activities, including school closures, starting on 23 March (Galizzi & Ghislandi 2020). Many other European countries followed soon, implementing similar public health measures. These measures were effective in curbing contagion rates in Europe, which slowed down over the summer months. Starting from September, European countries were confronted with a second wave of contagion, which led public authorities to opt for strict containment measures again. By 31 December 2020, the number of confirmed COVID-19 cases worldwide surpassed 80 million, of which over 26.1 million were recorded in Europe and over 2.1 million in Italy.[7] As of the same date, the toll of official COVID-19 deaths worldwide totalled to 1.82 million, to over 500,000 in Europe and to over 75,000 in Italy.

4.3 Data and Methods

In this section we present the data and the methodological approach used for estimating the mortality impact of the three flu pandemics under study in Italian provinces. Since the type of data we rely on to estimate the impact of flu pandemics in the 1950s-1960s differs from the type of data we have access to for estimating the impact of COVID-19, we adopt distinct methodologies as illustrated below.

[7] Source: WHO Covid Dashboard. See `https://covid19.who.int/`

4.3.1 Data and methods for the 1957-1958 and 1969-1970 pandemics

Official statistics on deaths and resident population at the provincial level are made digitally available by the Italian National Institute of Statistics (ISTAT) starting from 1982 only. We overcame this limitation by digitizing province-level census population data for years 1951, 1961, 1971 and 1981, disaggregated by sex and single-year age class (0, 1, ..., 100+), and yearly statistics on live births by sex and deaths by sex and five-year age class (0, 1-4, 5-9, ..., 95-99, 100+) at the provincial level for years 1953-1981. We additionally digitize regional level data on all-cause monthly deaths by sex for years 1953-1981. Since these data are not stratified neither by age nor by province, they do not allow to properly estimate mortality rates at the local level. Still, they are useful to get a sense of the timing of pandemic waves over years 1957-1958 and 1968-1970 in Italy and across regions.

We combine birth, mortality, and census population statistics to track yearly changes in the age structure of the provincial population and derive intercensal population estimates, for men and women separately. A detailed explanation of the methodological approach adopted to reconstruct province-level intercensal population for years 1953-1981 is available in Appendix D.1. We use inter-censal population estimates and mortality data to calculate yearly age- and sex-specific mortality rates for each province that are used to construct sex-specific provincial life tables for years 1953-1981 following standard procedures (Wachter 2014). Then, we estimate baseline sex-age-specific mortality rates that would have prevailed had the pandemics not occurred. We do so for each province and for Italy as a whole by omitting pandemic years (1957, 1958, 1968, 1969 and 1970) from the 1953-1981 annual series and by interpolating mortality rates in pandemic years through splines as in Viboud et al. (2013). We also omit year 1956 whose winter was characterized by a sharp increase in mortality due to an exceptional cold wave which hit Western Europe in February 1956 (Blangiardo 2020). We use baseline mortality rates to estimate baseline life expectancy and baseline lifespan variation at birth, measured in terms of standard deviation in age-at-death, in pandemic years. We calculate 95% confidence intervals for life expectancy and lifespan variation at birth using Monte Carlo simulation methods, assuming deaths count follow a binomial distribution (Andreev & Shkolnikov 2010; Chiang 1984).

We calculate excess mortality by sex and age class for each province and for Italy as a whole as the difference between baseline mortality and observed mortality rate in the pandemic years. When comparing excess mortality rates across provinces in pandemic years, province-level sex-specific excess mortality rates are age-standardised using the Italian population as standard. We further contrast baseline and observed life expectancy and lifespan variation, decomposing the difference by sex and age group for each province and for Italy as a whole in pandemic years through stepwise decomposition (Andreev & Shkolnikov 2012). The stepwise decomposition method can be used to decompose differences in any aggregate measure that depend only on the vectors of age-specific mortality rates. This approach allows us to assess which age groups have primarily contributed to the difference between baseline and observed life expectancy and lifespan variation in Italy in pandemic years.

4.3.2 Data and methods for the COVID-19 pandemic

Since the onset of the COVID-19 pandemic, the Italian National Institute of Statistics (ISTAT) has been releasing all-cause daily deaths count at the municipality level. At the time of writing,[8] such data cover the period between 1 January 2011 and 31 March 2021. We focus on calendar years 2011-2020 and compile daily deaths count for all causes at the municipality level, stratified by sex and five-year age classes (0, 1-4, 5-9,..., 95-99, 100+), between 1 January 2011 and 31 December 2020, aggregating them across provinces. We

[8] The current version of the chapter was finalised on June 15, 2021.

also compile ISTAT data on resident population at the municipality level, stratified by sex and single-year age classes on 1 January of years 2011-2020, and aggregate them across provinces. We reclassify population into five-year age groups so as to match the age classification used by ISTAT for deaths count and aggregate them at the provincial level. For years 2011-2019, we estimate mid-year population for each province-sex-age group combination as the average between population at the beginning and at the end of the year[9]. For year 2020, we estimate mid-year population for each province-sex-age group combination by subtracting half of the deaths occurring during 2020 from population at the beginning of 2020. In other words, mid-year population in 2020 is calculated as:

$$Pop^{mid}_{2020,x,s,p} = Pop^{1Jan}_{2020,x,s,p} - \frac{1}{2} Deaths_{2020,x,s,p}$$

where subscripts x, s and p indicate age, sex and province, respectively.

We exploit the daily nature of deaths count released by ISTAT to estimate the cumulative number of weekly excess deaths in 2020 since the beginning of the pandemic, for Italy as a whole and for each province separately. For this purpose, we first estimate the expected number of weekly deaths in the absence of COVID-19 (baseline) by fitting a generalised Poisson, Serfling-type linear model adjusted for year-to-year seasonality (Nielsen et al. 2018; Serfling 1953). This model is commonly used to estimate baseline mortality during influenza epidemics (Acosta et al. 2019; Andreasen et al. 2008). The basic structure of the model features trigonometric terms to account for seasonality and looks as follows:

[9]For each year $t \in (2011, 2019)$, we approximate the population on 31 December t with the population on 1 January *t+1*.

$$
\begin{aligned}
log(E(D_i)) = \beta_0 + \beta_1 time_i + \gamma_2 sin\Big(\frac{2\pi weekofyear_i}{52}\Big) + \gamma_3 cos\Big(\frac{2\pi weekofyear_i}{52}\Big) + \\
+ \gamma_4 sin\Big(\frac{2\pi weekofyear_i}{26}\Big) + \gamma_5 cos\Big(\frac{2\pi weekofyear_i}{26}\Big) + log(\psi_i)
\end{aligned}
\tag{4.1}
$$

where $E(D_i)$ is the expected number of deaths in a given week and population stratum i, and ψ_i is population exposure for stratum i. All the terms are further interacted with sex and age. To estimate population exposure (i.e. average weekly population), we use standard interpolation techniques (Dougherty et al. 1989). To make sure the model provides reliable estimate at the provincial level, we aggregate deaths count and population estimates over six age groups (0-14, 15-44, 45-64, 65-74, 75-84, 85+). We fit the model to the weekly deaths count starting from the week beginning on 3 January 2011 to the week beginning on 10 February 2020[10], and we project such baseline forward until week 53 of 2020 (starting on 28 December 2020). Excess deaths are then calculated as the difference between observed and expected deaths, for each population stratum of interest.

We further estimate life expectancy at birth and lifespan variation at birth by sex, for Italy as a whole and for each province, in years 2011-2020 using yearly deaths count and mid-year population estimates aggregated over five-year age classes (0, 1-4, 5-9, ..., 95-99, 100+) applying standard demographic techniques (Wachter 2014). We calculate 95% confidence intervals for the metrics of interest using Monte Carlo simulation methods, assuming deaths count follow a binomial distribution (Andreev & Shkolnikov 2010; Chiang 1984). Finally, we decompose the difference in observed life expectancy and lifespan inequality by age for each province and for Italy in 2020 compared to 2017-2019 average, through stepwise decomposition (Andreev & Shkolnikov 2012).

[10]The first case of local transmission of COVID-19 was registered in Italy on 21 February 2020.

4.4 Results

In this section, we illustrate the estimated mortality impact of the pandemic events under study in Italian provinces. A map of Italy highlighting provinces mentioned throughout Section 4.4 is shown in Figure D.1.

4.4.1 The mortality impact of the 1957-1958 and 1968-1970 pandemics in Italian provinces

Panel A of Figure 4.1 plots the ratio between observed monthly deaths in Italy between January 1957 and December 1958 with respect to the 1953-1955 average.[11] Trends are highly consistent with the diffusion patterns of the 1957-58 pandemic described in the Background section. Excess mortality is recorded from May 1957 throughout January 1958, the peak occurring visibly in October 1957 (30% above the 1953-1955 average). Regional-level trends suggest that both the timing and the mortality toll of the Asian flu were fairly homogeneous across Italian regions (Figure D.2 to D.4 in Appendix D). Focusing on the 1968-1970 flu pandemic, Panel B of Figure 4.1 plots the ratio between observed monthly deaths in Italy between January 1968 and December 1970 with respect to the 1965-1967 average. It is worth stressing that the mortality peak recorded in January 1968 (40% above the 1965-1967 average) cannot be attributed to the Hong Kong flu, as first infections were reported in Hong Kong in the summer of 1968. In fact, excess mortality recorded in early 1968 in Italy is ascribable to a recrudescence of the Asian flu virus activity, in its last antigenic variant (Ragona et al., 1978). The Hong Kong flu took most of its toll at the end of 1969, when mortality spiked 70% above the baseline. The mortality impact of the the Hong Kong flu was rather heterogeneous across Italian regions. Indeed, while some Northern regions, like Trentino Alto-Adige, reported >100% higher mortality compared to the baseline in December 1969, other regions in the South,

[11] As said, we exclude 1956 from the baseline computation due to high mortality levels recorded in February 1956, when Italy was hit by an exceptionally cold wave, which caused about 50,000 excess deaths (Blangiardo 2020).

like Sicily, were markedly less affected (Figure D.5 to D.7 in Appendix D).

Panel A of Figure 4.2 plots the evolution of life expectancy at birth in Italy in 1953-1982, for men and women separately, against the estimated baseline, along with the respective 95% confidence intervals (shaded areas).[12] Statistically significant negative deviations from baseline life expectancy emerge around pandemic years 1957-1958 and 1969, for both men and women.[13] For men, observed life expectancy at birth in 1957, 1958 and 1969 for Italy as a whole is 1.3, 0.4 and 0.7 years lower than the baseline, respectively. Such negative deviations from baseline life expectancy reflect increased mortality rates. We estimate 9.5 excess deaths per 10,000 men in 1957, 2.2 in 1958 and 6.7 in 1969. In the same years, we find female life expectancy at birth to be 1.1, 0.3 and 0.6 years lower than expected. Estimated excess deaths per 10,000 women are 7.6 in 1957, 1.7 in 1958 and 3.8 in 1969. Overall, we estimate a total of about 42,000 and 28,000 excess deaths in peak years 1957 and 1969, respectively. As for lifespan variation, deviations from the baseline in pandemic years at the national level are hardly remarkable, except for a slight increase around 1957-1958 (Figure 4.2, Panel B). Estimates of observed and baseline life expectancy and lifespan variation at birth as well as excess deaths for Italy as a whole in pandemic years are reported in Table 4.1.

Turning to province-level analysis, we find the mortality burden of the 1957-1958 pandemic to be homogeneously spread over the Italian peninsula, the effect being most clearly felt in the first year of the pandemic. Some clustering emerges in the Centre-South around densely populated areas, such as the provinces of Rome and Naples in the Centre-South, and the province of Turin in the North-West. Simple regression analysis confirms that excess mortality in peak year 1957 correlates positively with population density at the

[12]Since estimates are very precise, shaded areas are hardly visible.

[13]Deviations from baseline life expectancy recorded in 1956 and in 1968 are attributable to events other than the pandemic waves under study (see Background section).

provincial level, after accounting for the share of employment in industry and for region fixed effects (Table 4.2, Column 2). In 1957, in the most hardly hit areas, all-cause excess mortality amounts up to 30 per 10,000 people, for both men and women (Figure D.8), while life expectancy at birth is 1.5 to 2 years lower than expected (Figure 4.3). In provinces were life expectancy contracts more sharply, lifespan variation tends to be higher than expected, up to 1.5 years, both in 1957 and 1958 (Figure 4.4 and Figure 4.5).

The analysis relative to the 1969-1970 pandemic wave delivers a somewhat different picture.[14] First of all, we document a geographical gradient in mortality, Northern provinces being more affected than Southern ones, especially in the first year of the pandemic. In 1969 most provinces in the North display all-cause excess mortality, with peaks of up to 30 excess deaths per 10,000 people among men in the North-East (Figure D.9). Several Southern provinces, instead, display lower-than-expected mortality in both pandemic years. Second, we estimate the impact of the 1969-1970 pandemic on life expectancy at birth to be larger compared to the impact of the 1957-1958 pandemic in the hardest hit areas, particularly in the case of men. In 1969, we estimate life expectancy at birth to be up to 3.1 years lower than the baseline in the case of men, and up to 2 years lower in the case of women in most affected provinces (Figure 4.6). Thus, while the 1957-1958 pandemic had a larger impact on mortality for Italy as whole, the death toll of the 1969-1970 pandemic was more geographically clustered, hitting more severely in specific areas, mostly in the industrialized provinces of the North. Simple regression analysis suggests that excess mortality in 1969 concentrated, indeed, in industrial areas (Table 4.2, Column 3). Third, we find the impact of the 1969-1970 pandemic on lifespan inequality to be inversely associated to the impact on life expectancy, albeit less clearly than in the case of the 1957-1958 pandemic (Figure 4.7 and Figure 4.8).

[14]We focus on 1969 and 1970 as the pandemic hit Italy in late 1969 (see Background section).

To gain further insights about such differences, we investigate patterns in age- and sex-specific excess mortality at the provincial level in pandemic years 1957 and 1969, when mortality peaks occurred. In 1957, excess mortality concentrates among individuals aged 75+ in the case of men and among individuals aged 80+ in the case of women (Figure 4.9). In 1969, instead, excess mortality is strongly concentrated among relatively younger individuals, aged 60-79, for both men and women (Figure 4.10). It is worth noticing, though, that in both pandemic events children, particularly those aged <1, were not spared by excess mortality either. In fact, age-based decomposition of the difference between baseline and observed life expectancy in pandemic years suggests that increased infant mortality rates contribute substantially to negative deviations from the baseline (Figure 4.11). This is particularly evident in hardest hit provinces, such as Turin (North-West) in 1957 and Trento (North-East) in 1969. For instance, higher-than-expected mortality among children aged <1 in Turin in 1957 contributes to about -0.5 out of -2.3 years in the difference between observed and baseline life expectancy in the case of men, and to about -0.2 out of -1.5 years in the case of women. The increase in lifespan variation observed in hardest hit provinces in both pandemic events is also driven by increased infant mortality rates. Indeed, higher mortality rates among children aged 0-5 imply an expansion in lifespan variability which more than compensates the mortality compression implied by increased mortality among the elderly, particularly in 1957 (Figure 4.12).

4.4.2 The provisional mortality burden of the COVID-19 pandemic

Panel A of Figure 4.13 plots the ratio between observed weekly deaths in Italy throughout 2020 with respect to the baseline by sex, across all ages. Two waves are clearly discernible in 2020: the first hit in early March, with a peak in week 12 when male mortality spiked nearly 200% above the baseline, and the second one hit in mid-October. No excess mortality was recorded over the summer months. Such patterns are highly consistent with the timing of containment measures adopted by local authorities as briefly outlined in the Background section. It is worth noting that we estimate no mortality deficit over the summer months, either. This suggests the absence of short-term harvesting dynamics (Toulemon & Barbieri 2005). Indeed, had the first wave of COVID-19 killed predominantly the most fragile individuals, we should have observed a decline in mortality compared to the baseline over the summer months. Panel B of Figure 4.13 plots the cumulative number of weekly excess deaths over the 2020 pandemic period. Overall, we estimate a total of about 113,600 excess deaths in Italy over weeks 8-53 of 2020 (59,600 men and 54,000 women), which is nearly 50% higher than the number of official COVID-19 deaths reported by the integrated surveillance system of the Italian National Institute of Health (*Istituto Superiore di Sanità*) over the same period (≈ 75,900).[15] Our estimate of total excess mortality in calendar year 2020 ascribable to COVID-19 is higher than the estimate provided by the Italian National Institute of Statistics (2021) (≈99,000 excess deaths). However, the latter represents a conservative estimate as it is based on a simple comparison between mortality levels in 2019 and 2020 which does not account for secular and seasonal trends. In any case, the sizeable discrepancy between all-cause excess mortality and COVID-19 related mortality over the pandemic period highlights the usefulness of taking an all-cause excess mortality approach to evaluate the mortality burden of epidemic/pandemic events. From Figure 4.14, it is evident that excess mortality inflates with age. We estimate no excess mortality for age groups 0-14 and 15-44, while excess deaths of age groups 45-64 and 65-74 represent about 6% and 13.5% of the total, respectively. The mortality toll increases substantially for the 75-84 (33% of the total) and for the 85+ (47% of the total). When considering Italy as a whole, life expectancy at birth in 2020 drops by about 1.4 and 1.1 years compared to 2019, for men and women respectively, landing back to the levels of the early 2010s (Figure 4.15, panel

[15] The number of deaths related to COVID-19 is reported on a daily basis by the Civil Protection Department of the Italian Government. Data are freely accessible at `https://github.com/pcm-dpc/COVID-19`.

A). Lifespan variation contracts too, by about 0.5 and 0.4 years for men and women respectively, compared to 2019 levels (Figure 4.15, panel B).

We now turn to province level analysis. Figure 4.16 plots all-cause excess mortality rate (per 10,000) in 2020 in each province, for men and women separately. It is evident that country-level estimates mask substantial heterogeneities. Excess mortality in 2020 concentrates sharply in provinces located in the North-West of the country, notably in the Lombardy region, while the toll in most provinces in the Centre-South is relatively modest. We find excess mortality rate to be highest in the provinces of Bergamo (73.6 for men, 53.9 for women), Brescia (45.9 men, 38.8 women), Cremona (68.2 men; 54.1 women), Lodi (59.9 men, 49.4 women), and Piacenza (52.2 men, 41.1 women), which were the epicentres of the first pandemic wave in March 2020. To evaluate the impact of increased mortality on life expectancy at the provincial level, in Panel A of Figure 4.17 we plot the difference between life expectancy at birth in 2020 vs the 2017-2019 average in each Italian province, for men and women separately. In line with excess mortality rates, life expectancy contracts visibly more sharply in the epicenters of the first pandemic wave, where life expectancy in 2020 is estimated to be 2.7 to 4.6 years lower compared to the 2017-2019 average in the case of men, and 2.2 to 3.0 years lower in the case of women. While some significant contraction in life expectancy at birth is observed also in a few provinces in the Centre-East and in the South-East, the estimated impact on life expectancy in the Centre-South of the country is markedly milder compared to the North, when not nil. Compared to the 2017-2019 average, lifespan variation in 2020 contracts in most provinces. Still, there appears to be no correlation between the impact on life expectancy and the impact on lifespan variation in the case of men, while the correlation is mildly negative in the case of women (Figure 4.18).

To fully characterize the unequal mortality impact of COVID-19, we explore excess mortality patterns by age and sex in each province. Excess mortality materializes among the 55+ and among the 65+ for men and women respectively, increasing linearly with age (Figure 4.19). When decomposing the age-specific contribution to changes in life expectancy and lifespan variation in 2020 with respect to the 2017-2019 average (Figure 4.20), it is clear that older populations play a major role. The drop in both life expectancy and lifespan variation is driven by increased mortality at 55-84 and 65-89, for men and women respectively. It worth noting that increased mortality among the 45-69 contributes *positively* to the change in lifespan variation Still, such expansion in lifespan variability is more than compensated by the mortality compression implied by increased mortality among the 70+.

4.5 Discussion

The disruptiveness of the COVID-19 pandemic has spurred global research efforts on its health, demographic and socio-economic consequences. To better seize its epochal effects, though, historical comparisons with past pandemics are needed. While the 1918 influenza pandemic has been extensively studied, little attention has been devoted to other influenza pandemics which caused important human losses worldwide between the 1950s and the 1960s. To the best of our knowledge, this is the first study to generate estimates for the total mortality burden of both the 1957-1958 and the 1968-1970 flu pandemics at the subnational level in a country, i.e. Italy, hardly hit by the COVID-19 pandemic, drawing a historical comparison with the latter.

Before discussing the main results of our study, we highlight its strengths and limits. A major caveat relates to the different types of mortality data we rely on in our analysis, which limit the comparability of the pandemic events under study. Indeed, while the empirical analysis for the COVID-19 pandemic is partly based on daily deaths count, the empirical analysis for the pandemics of the 1950s and of the 1960s is based on annual deaths count only. Studies on influenza-related mortality generally employ weekly or monthly mortality data, stratified by sex and age, to estimate excess mortality. However, this kind of data are not available for Italian provinces for pandemic years of interest in the 1950s-1970s. For this reason, we adopted an annualized approach, and estimated excess mortality in pandemic years of decades 1950s-1960s with respect to a model-predicted spline baseline, fitted to the historical series of mortality rates outside the pandemic periods, for each Italian province. A similar approach was used by previous studies to assess the global mortality burden of past influenza pandemics (Murray et al. 2006; Viboud et al. 2013). These works confirmed the general validity of the annualized approach to estimate total excess mortality. Turning to strengths, the subnational focus is a major advantage of our analysis since it allows to account for geographically clustered patterns in contagion, which is key to fully characterize the demographic consequences of pandemics. Methodologically, the all-cause excess mortality approach is a further strength of our work since it allows to assess both the direct and indirect human costs of pandemic events. Moreover, by quantifying total excess mortality in pandemic years, we are able to evaluate the overall impact on population well-being using metrics such as life expectancy and lifespan variation which enable effective and reliable comparison of major demographic shocks over time and place. It is important to keep in mind that life expectancy and lifespan variation are based on mortality rates observed during a given period. As such, they provide an estimate of the average life span, and of the surrounding uncertainty, of a group of individuals living under the observed mortality regime. However, mortality is rarely constant over time and, therefore, metrics based on period mortality rates, such as life expectancy, are usually poor indicators for the life span of an actual group of individuals (Luy et al. 2020). This holds true particularly for sudden shocks, such as epidemics, wars or natural disasters, which yield sizeable but temporary changes in mortality rates. Nevertheless, although no individual can be expected to experience the mortality regime of pandemic periods over the entire life course, life expectancy and lifespan variation are powerful tools for summarizing and comparing mortality shocks over regions and time, especially because they do not depend on populations' age structure (Marois et al. 2020).

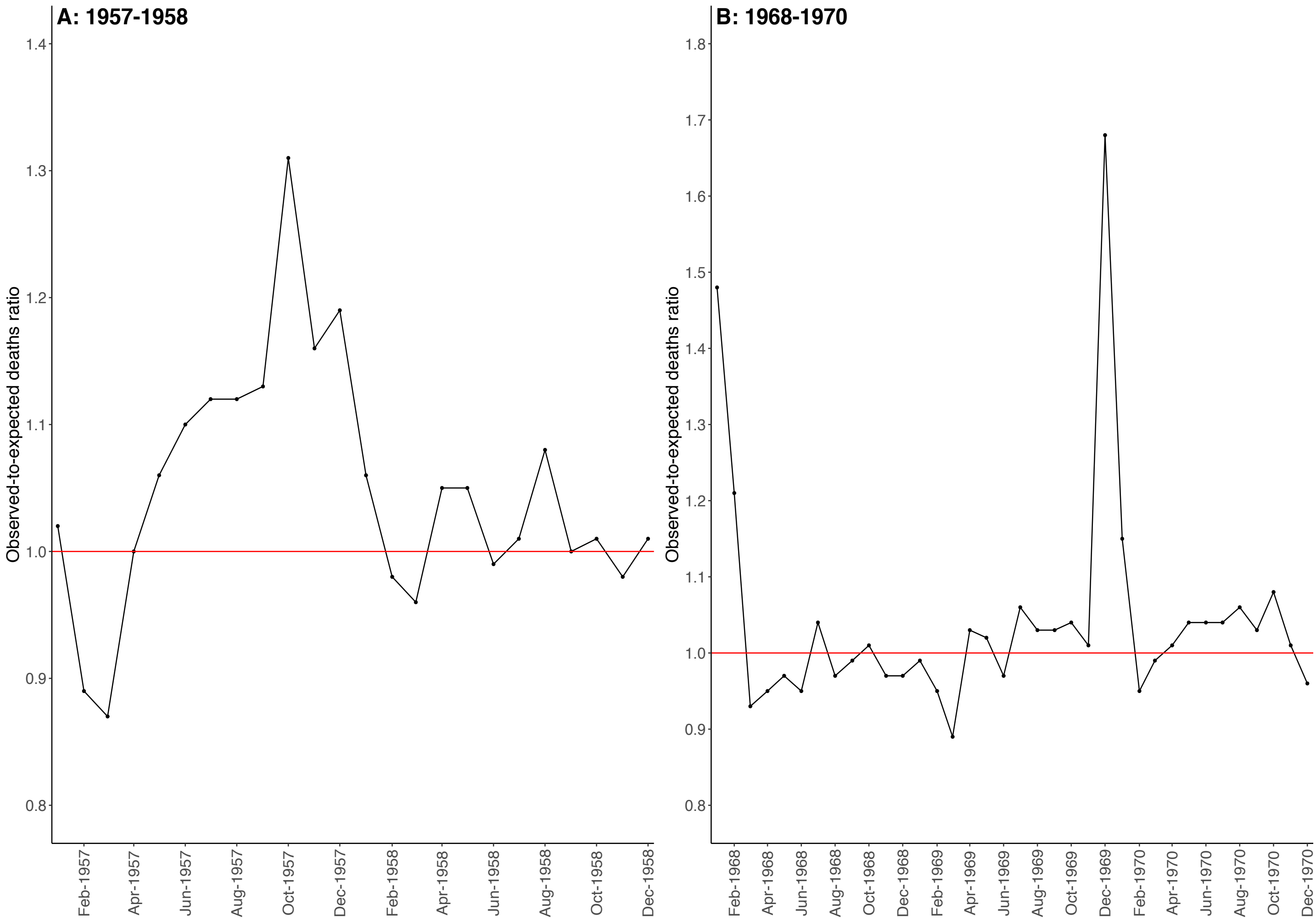
Our analysis suggests that the human costs of the 1957-1958 pandemic were overall larger than those of the 1969-1970 pandemic for Italy as a whole. However, the latter hit harder in specific provinces. We estimate that in peak year 1957, life expectancy for Italy as a whole was 1.3 and 1.1 lower than expected, for men and women respectively. In most severely hit provinces, observed life expectancy was 1.5 to 2 years lower compared to the baseline, for both sexes. In peak year 1969, instead, life expectancy at the country level was 0.7 lower than expected in the case of men, and 0.6 lower in the case of women. However, in most severely affected provinces, including most of the industrialized areas in the North of the country, the difference between observed and baseline life expectancy in 1969 was nearly up to five times larger than the difference recorded for Italy as a whole. Indeed, we estimate life expectancy in 1969 to be up to 3.1 and up to 2.1 years lower than expected in hardest hit areas, for men and women respectively. These results confirm the importance of subnational level analysis for grasping the true scale of pandemic severity. The sharper local impact on life expectancy of the 1969 pandemic wave, compared to the 1957 one, is explained by increased mortality rates among relatively younger age groups. In most severely hit areas, we find absolute excess mortality rates in 1957 to be lowest among school-age children and highest among the 75+, documenting a monotonic increase with age. In contrast, in 1969 excess mortality concentrates in the 60-79 age group, older people being largely spared in most Italian provinces. As documented through decomposition analysis, mortality among the 80+ does not contribute to the difference between observed and baseline life expectancy in 1969. These patterns are consistent with previous works stressing how 'senior sparing', i.e. lack of excess mortality at older ages, did not play a relevant role in the 1957-1958 pandemic, while it did so in the 1968-1970 pandemic (Viboud et al. 2013). In 1968-1970, just as in 1918, seniors were largely spared by pandemic-induced excess mortality due to protection granted by prior childhood exposure to antigenically related pathogens (Andreasen et al. 2008; Chowell et al. 2014; Olson et al. 2005; Viboud et al. 2013). In contrast, excess mortality was recorded in both pandemic events also among pre-school children, particularly those aged <1. Excess infant mortality, which has an expanding effect on the dispersion of age-at-death, compensating the compression induced by excess mortality among the elderly, explains also the higher-than-expected lifespan variation observed in hardest hit provinces in both 1957 and 1969.

When comparing the mortality impact of flu pandemics of the 1950s-1960s with that of COVID-19, a few similarities emerge. In particular, high population density areas, such as the industrialized provinces of Northern Italy, appear as particularly vulnerable to pandemic events. Indeed, despite differences in the health care system and in the population structure, the North-South mortality gradient documented for the 1969 pandemic wave is suggestively analogous to the provisional mortality pattern of the COVID-19 pandemic. Moreover, both the pandemic episodes of the 1950s-1960s and the COVID-19 pandemic are characterized by a gender gradient in mortality, human costs being systematically higher among men as compared to women. As stressed by previous works, explanations may include biological as well as social and behavioural factors (Galasso et al. 2020; Krieger et al. 2020). Despite these similarities, several differences between the COVID-19 and the 1950s-1960s pandemics stand clearly out. First of all, mortality associated to the COVID-19 pandemic in Italian provinces is more spatially clustered. Excess mortality in 2020 concentrates sharply in Northern provinces, especially in the Lombardy region, where first clusters of COVID-19 were detected (Blangiardo et al. 2020). Such geographical clustering is ascribable to strong containment measures, adopted starting from early March 2020, which prevented the virus from spreading uncontrolled throughout the country (Gatto et al. 2020). Second, in the absence of tight containment measures, the pandemics of the 1950s and 1960s took their toll over very few months. In the case of the Hong Kong flu, excess mortality was recorded between December 1969 and January 1970. This is not the case of the COVID-19 pandemic, which has been unfolding over multiple waves whose timing is consistent with that of containment policies implemented by public health authorities. Third, while pandemic waves of the 1950s and the 1960s claimed lives at both extremes of the age spectrum, the COVID-19 has taken its toll mostly among individuals aged 65+, children and young adults (< 45) being spared by excess mortality. Fourth, notwithstanding the implementation of tight containment measures, the strong age gradient in excess mortality, and the availability of more advanced medical technologies, human costs of the COVID-19 pandemic are substantially heavier than those of pandemic episodes of the 1950s and 1960s. In the Italian epicentres of the first wave of the COVID-19 pandemic, life expectancy at birth in 2020 is estimated to drop as much as by 4.6 years for men and 3 years for women compared to the 2017-2019 average, back to the levels recorded in the early 2000s. Fifth, contrarily to the 1957-1958 and 1969-1970 pandemic episodes, in 2020 lifespan variation moved in the same direction of life expectancy, i.e. it decreased. In a context where deaths occur mostly at older ages, the sharp increase in mortality rates among seniors due to the COVID-19 pandemic increases average age-at-death while reducing its variation. Such positive correlation between these two metrics in the aftermath of the COVID-19 pandemic in Italy is analogous to patterns in life expectancy and lifespan variation detected in England and Wales based on all-cause mortality data for 2020 (Aburto, Kashyap, et al. 2021).

In conclusion, while documenting that the mortality impact of the flu pandemics of the 1950s and the 1960s in Italy was sizeable, our analysis shows that in the hardest hit areas of the country the COVID-19 pandemic, whose trail of death is still unfolding at the time of writing, represents the major mortality shock since the end of World War II.

Figures Chapter 4

Figure 4.1: Ratio between observed and baseline all-cause monthly deaths in Italy 1957-1958 and 1968-1970



Notes: For each month of 1957-1958, the baseline is computed as the 1953-1956 average of total deaths. For each month of 1968-1970, the baseline is computed as the 1965-1967 average of total deaths. Own elaboration based on ISTAT data.

Figure 4.2: Observed vs baseline life expectancy and lifespan variation at birth – Italy 1953-1982

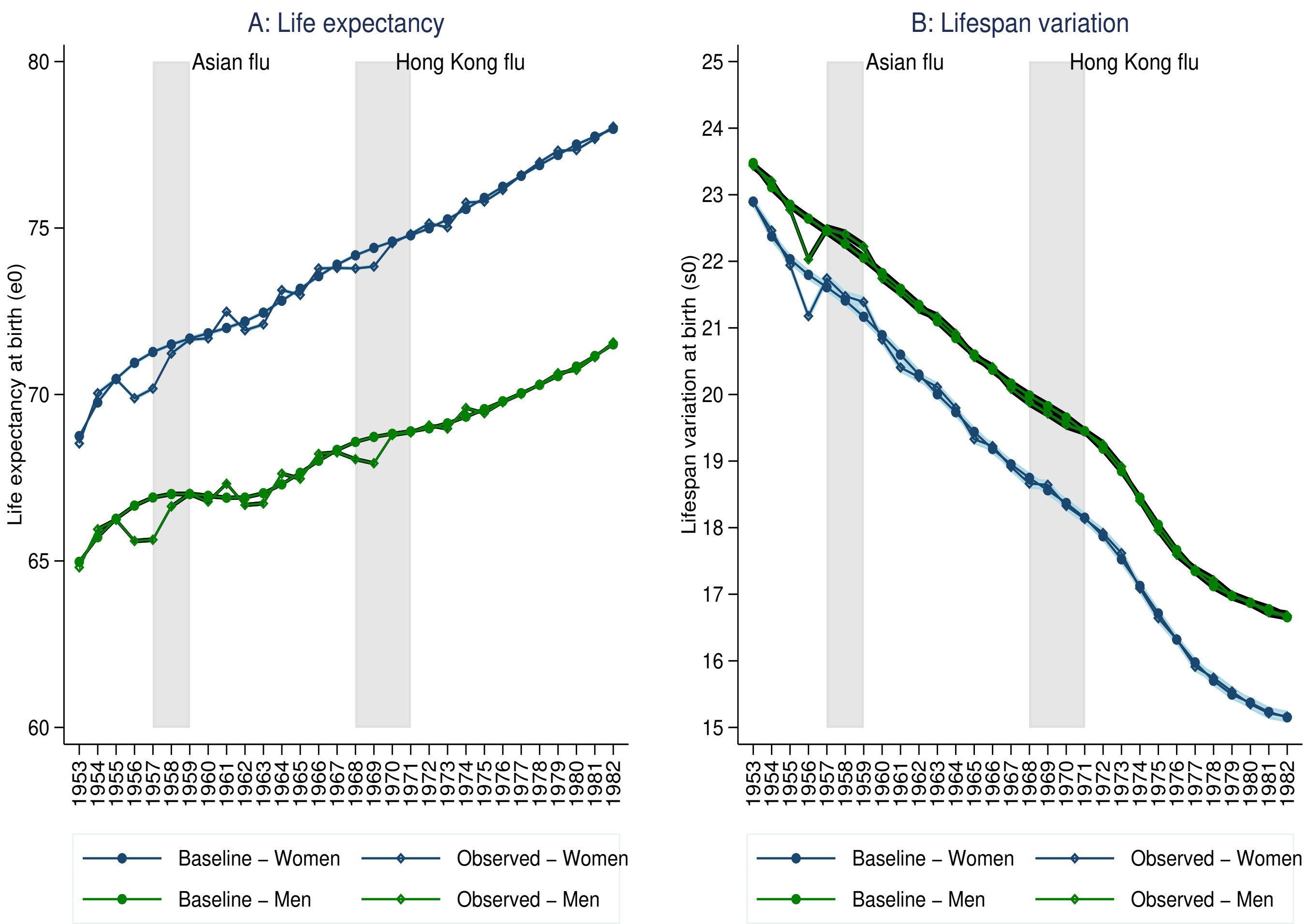


Figure 4.3: Observed vs expected life expectancy at birth in Italian provinces 1957-1958

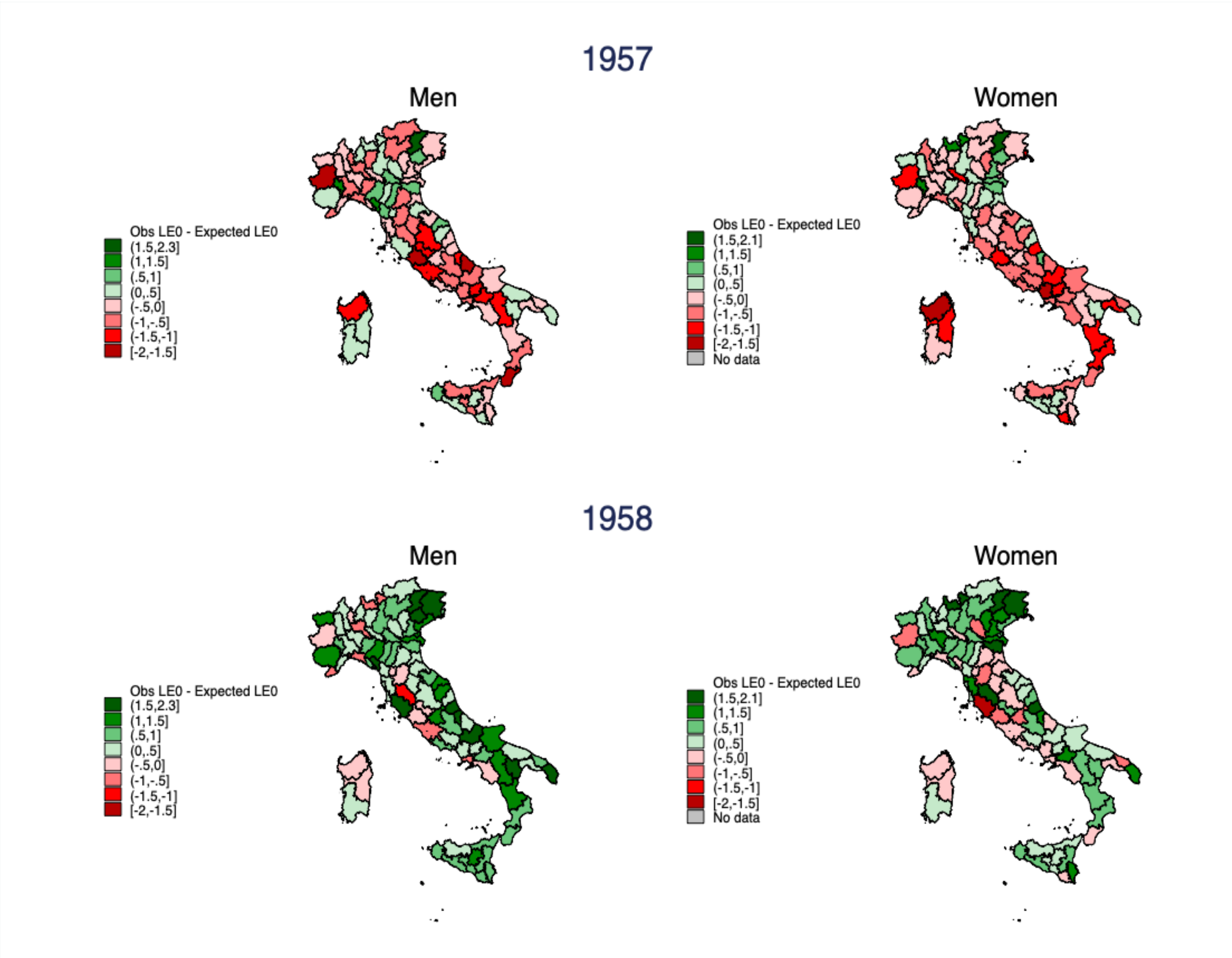


Figure 4.4: Observed vs expected lifespan variation at birth (standard deviation) in Italian provinces 1957-1958

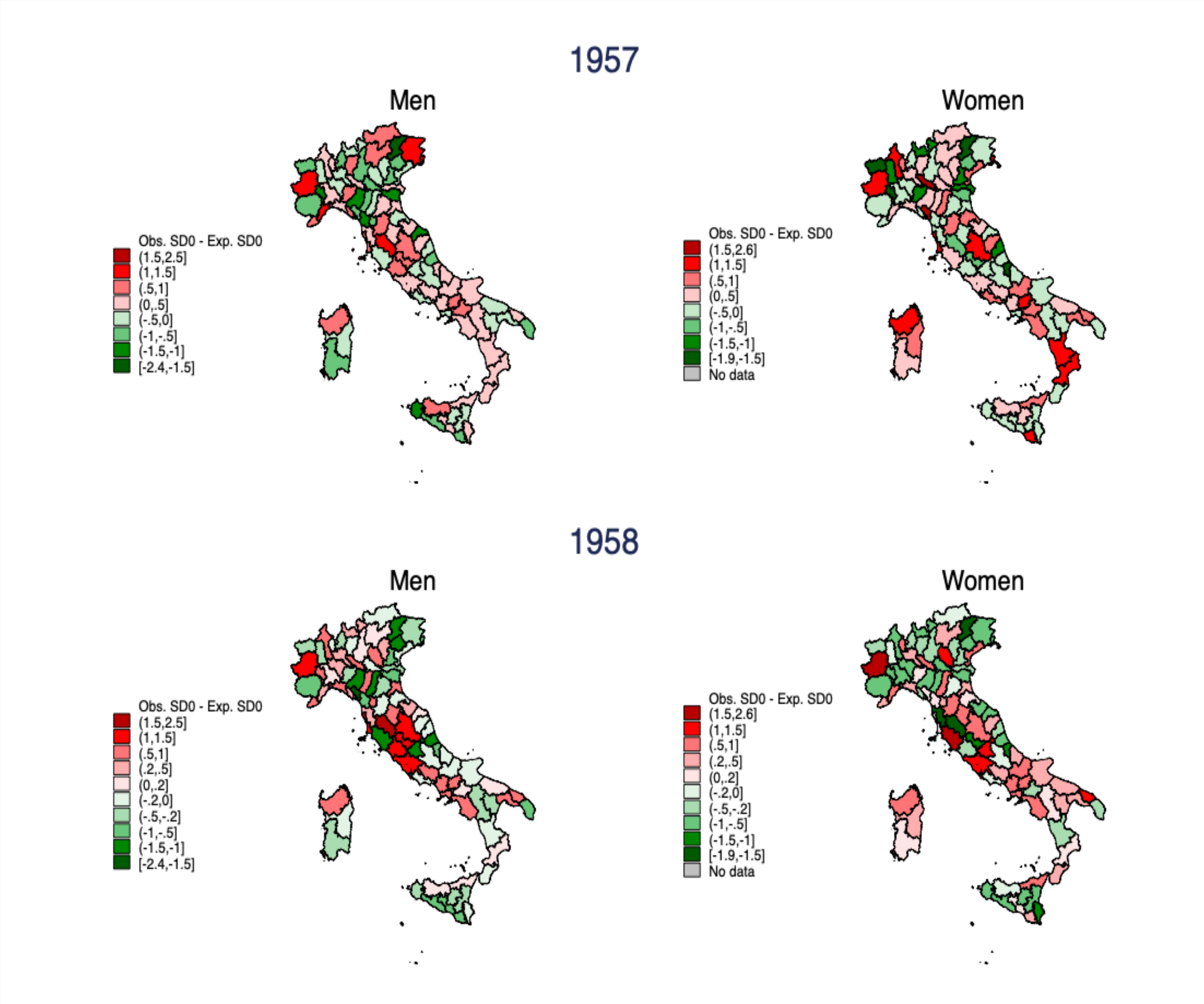


Figure 4.5: Impact on lifespan inequality at birth vs impact on life expectancy at birth in Italian provinces 1957-1958

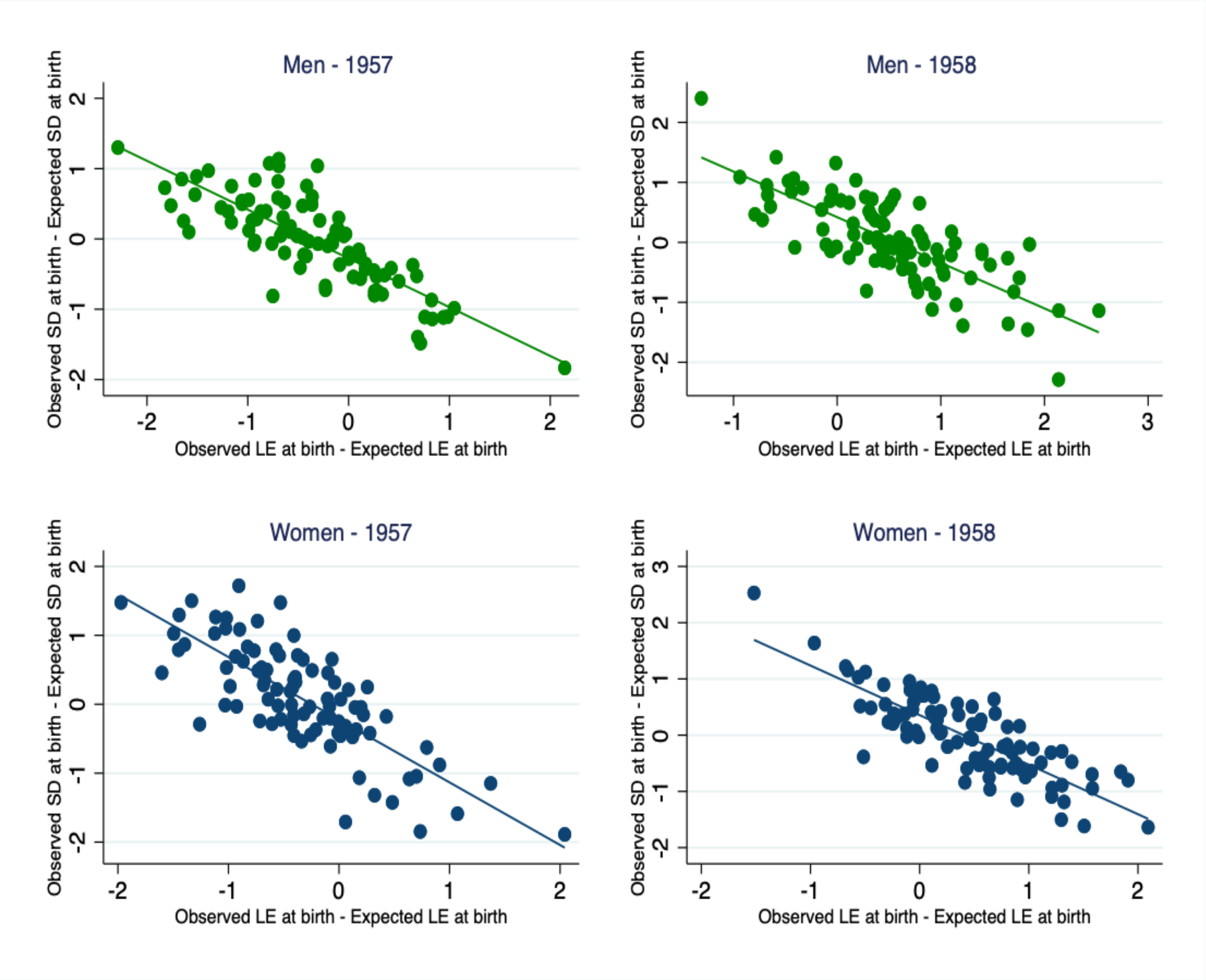


Figure 4.6: Observed vs expected life expectancy at birth in Italian provinces 1969-1970

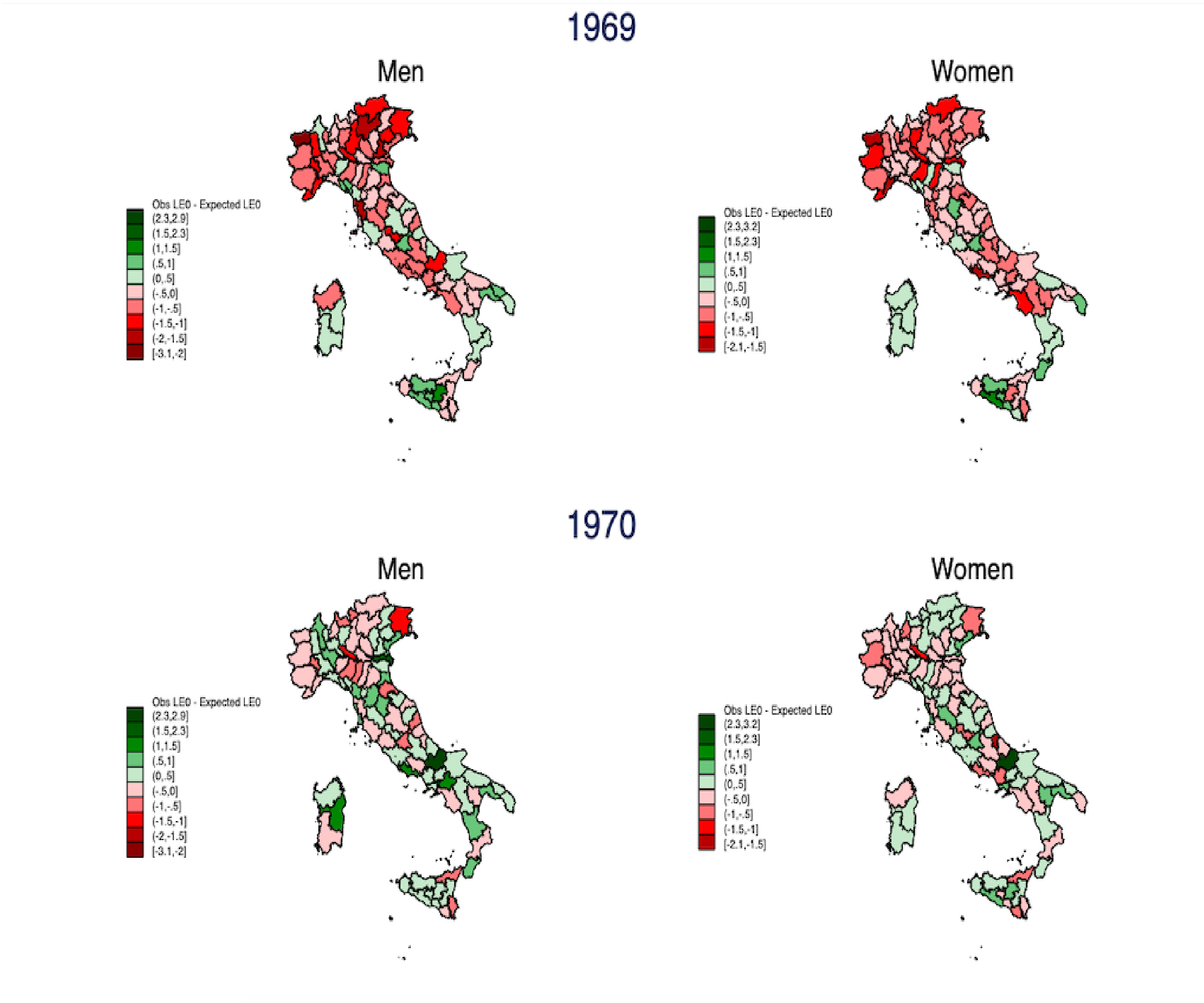


Figure 4.7: Observed vs expected lifespan variation at birth (standard deviation) in Italian provinces 1957-1958

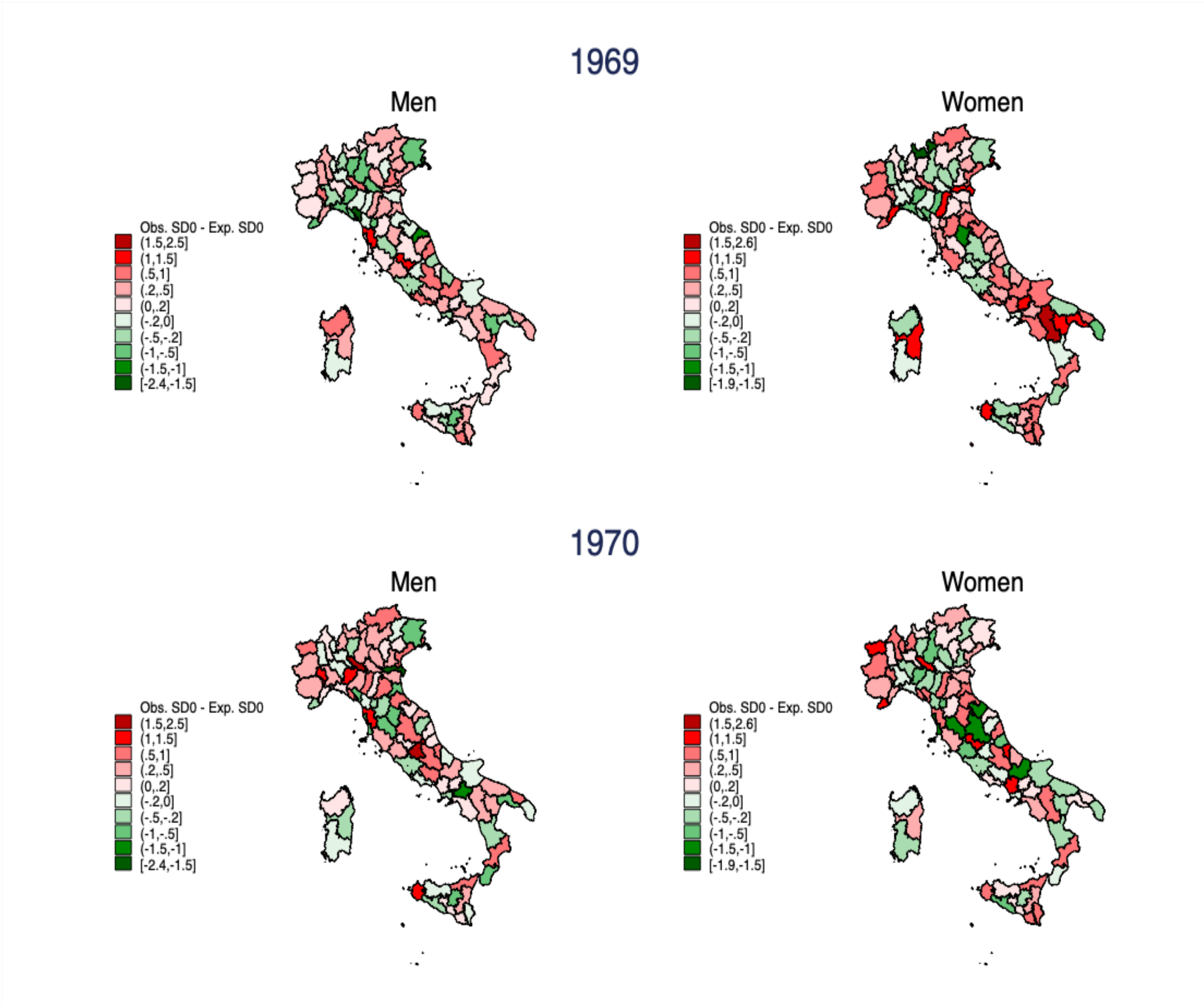


Figure 4.8: Impact on lifespan inequality at birth vs impact on life expectancy at birth in Italian provinces
1969-1970

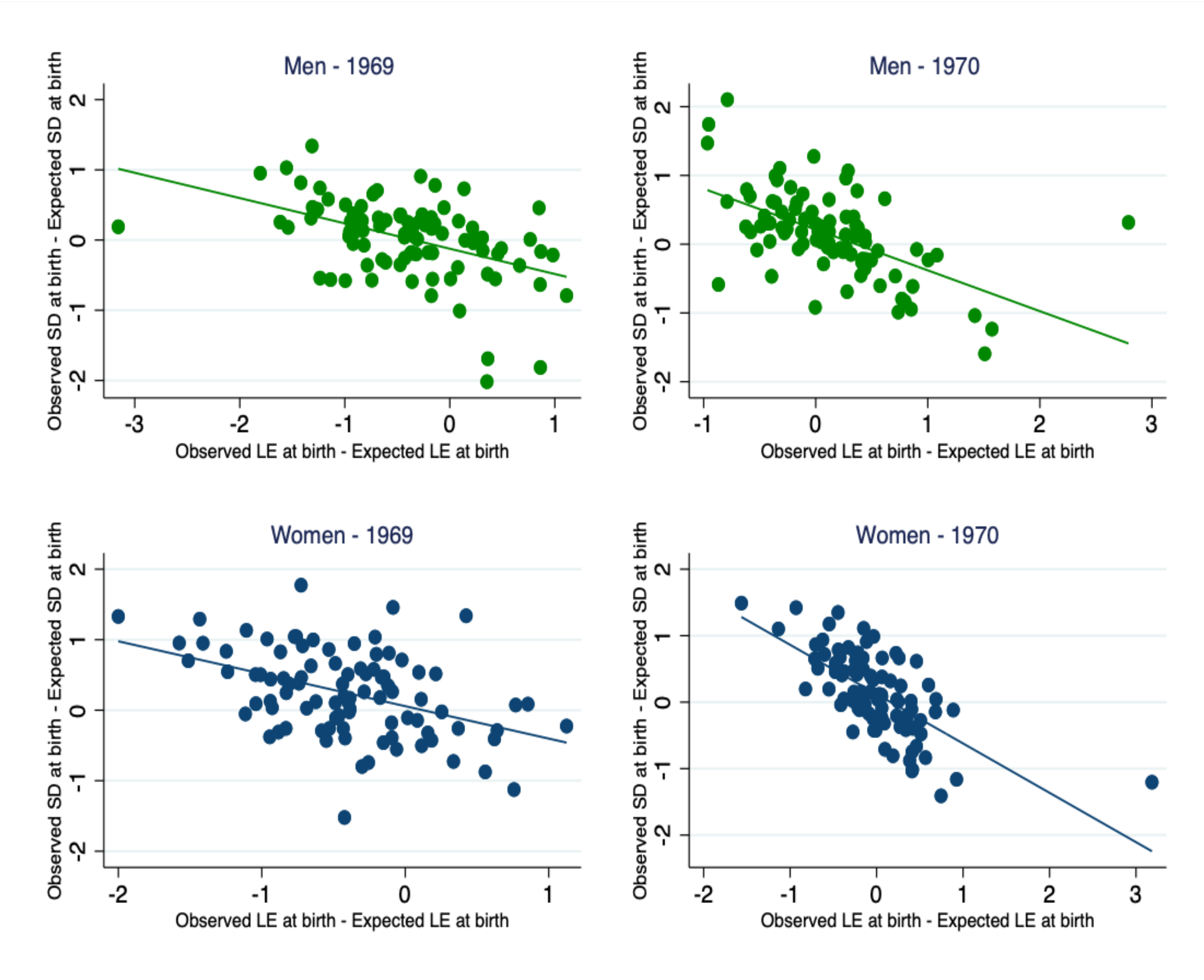
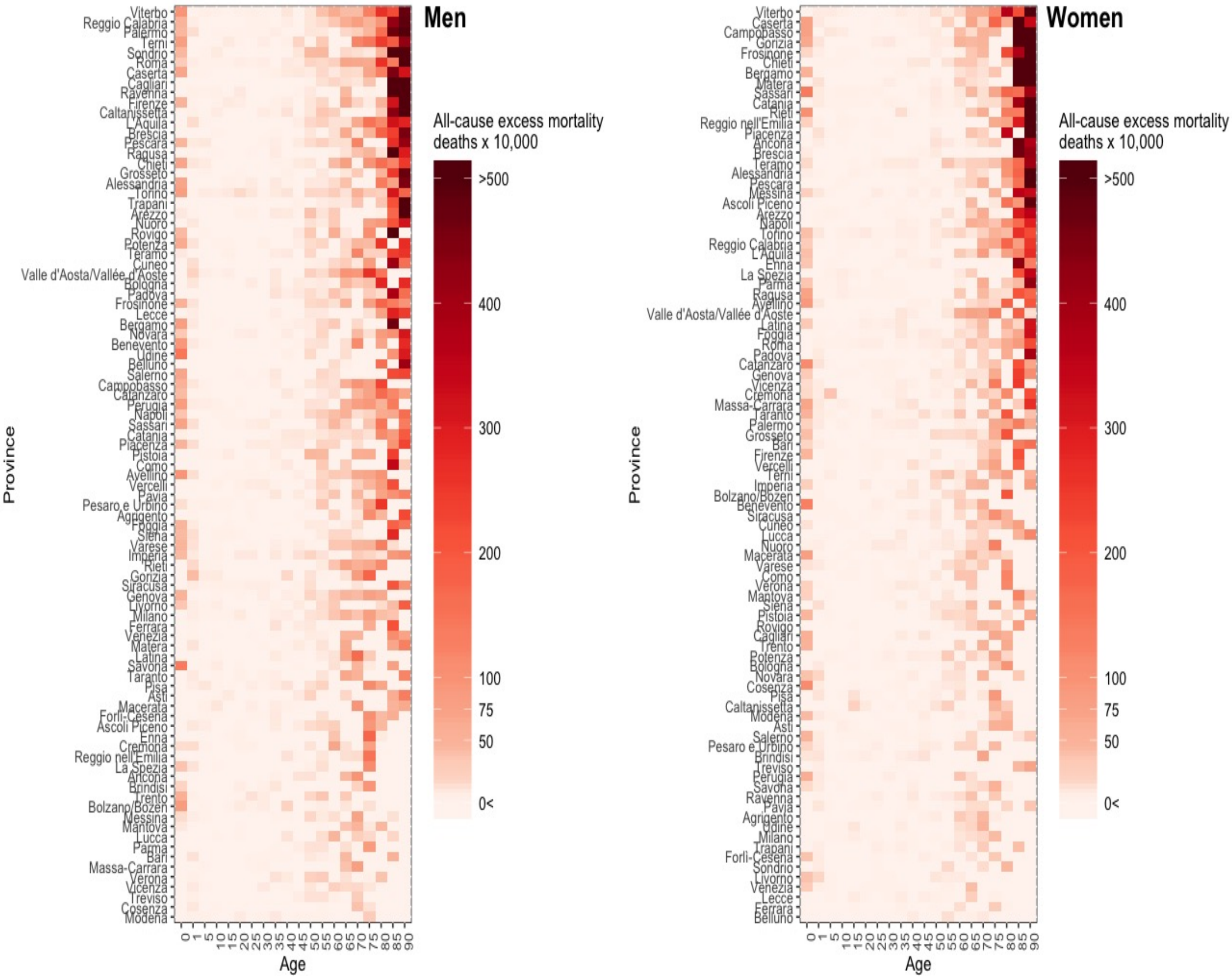
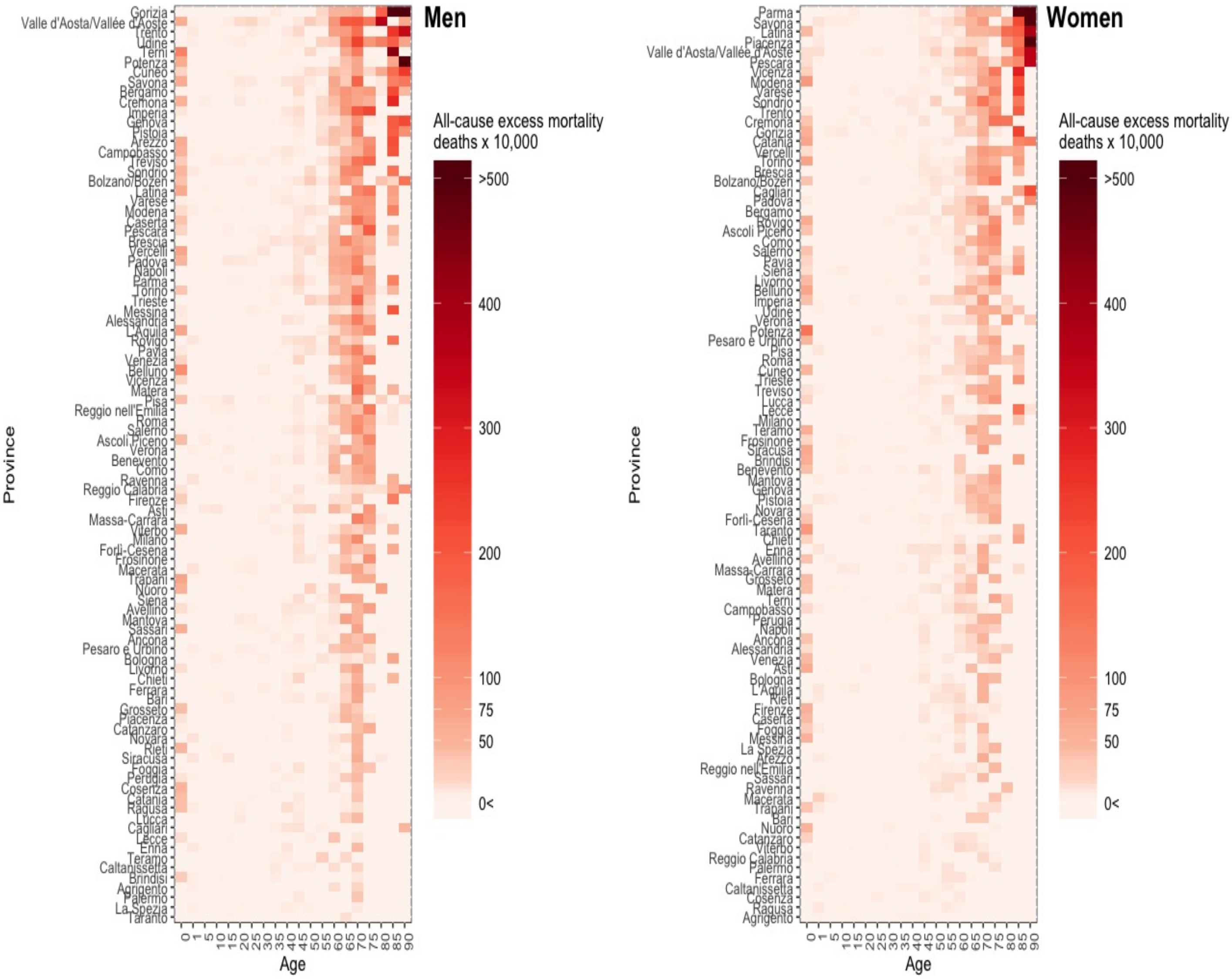


Figure 4.9: All-cause excess mortality by age and province (deaths x 10,000) 1957



Notes. Excess mortality rate calculated as the difference between observed and expected mortality rate. Provinces are ordered by (total) age-standardized excess mortality rate.

Figure 4.10: All-cause excess mortality by age and province (deaths x 10,000) 1969



Notes. Excess mortality rate calculated as the difference between observed and expected mortality rate. Provinces are ordered by (total) age-standardized excess mortality rate.

Figure 4.11: Age-specific contribution to difference between observed and baseline life expectancy at birth 1957 & 1959

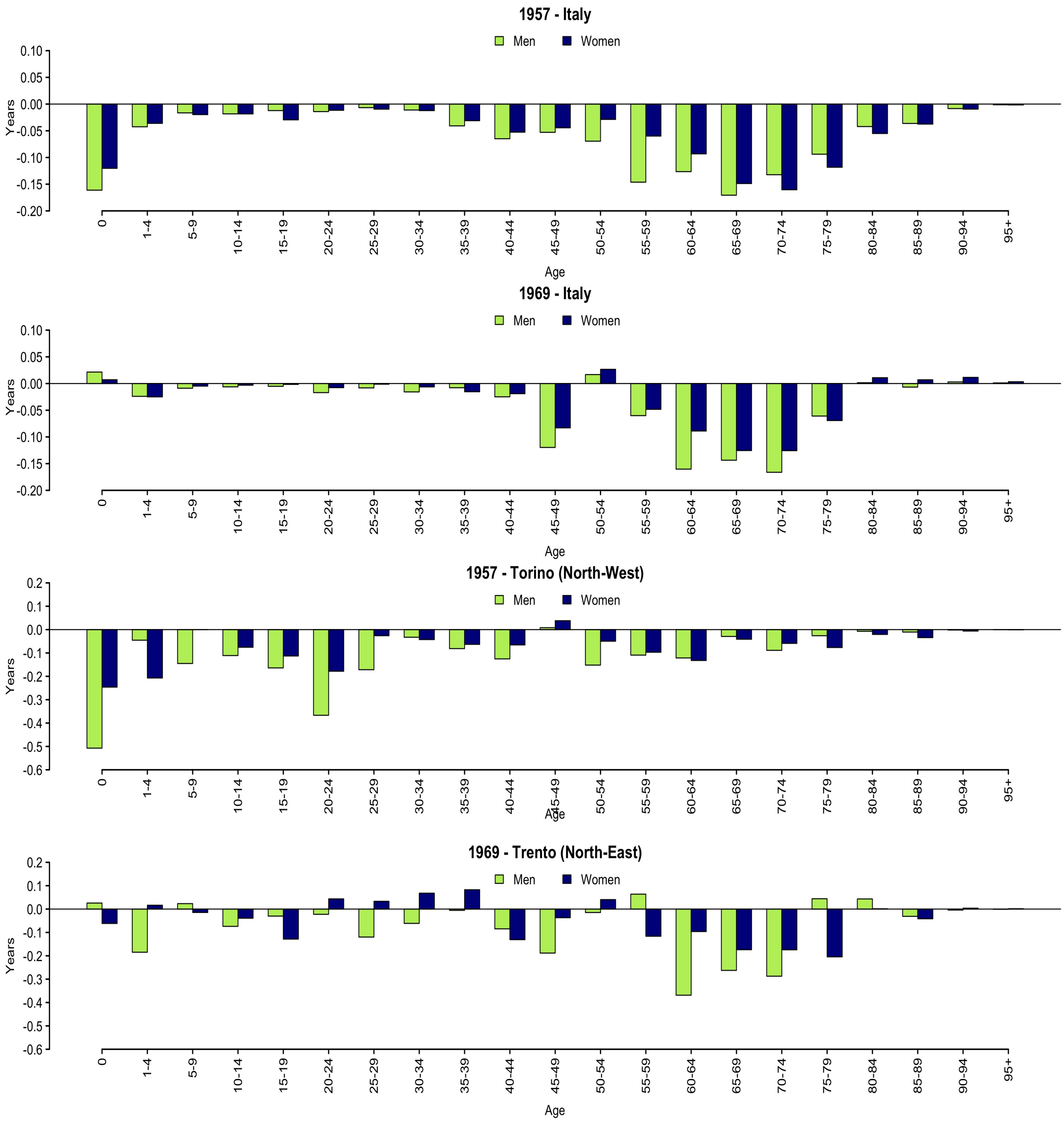


Figure 4.12: Age-specific contribution to difference between observed and baseline lifespan variation at birth 1957 & 1969

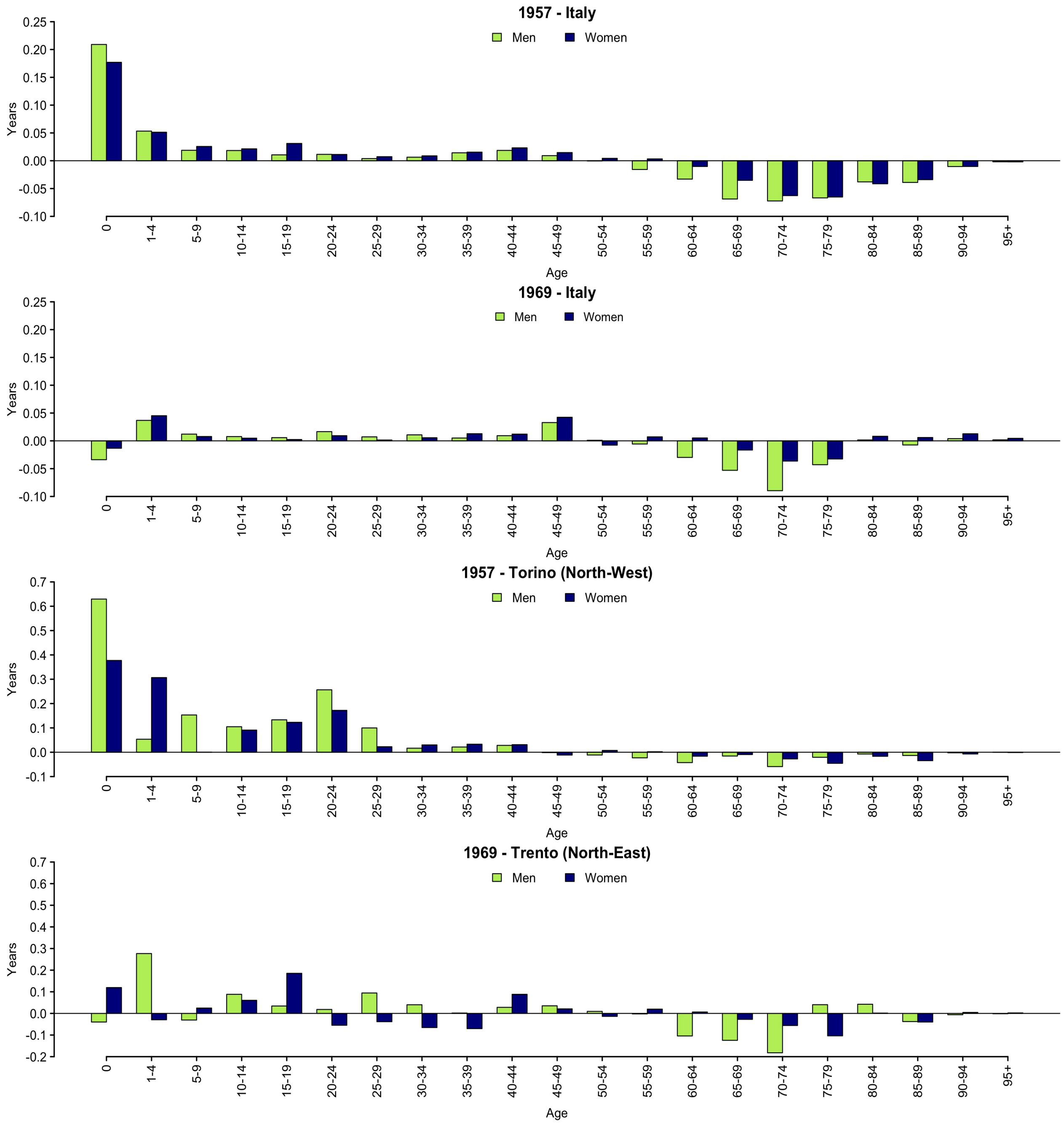
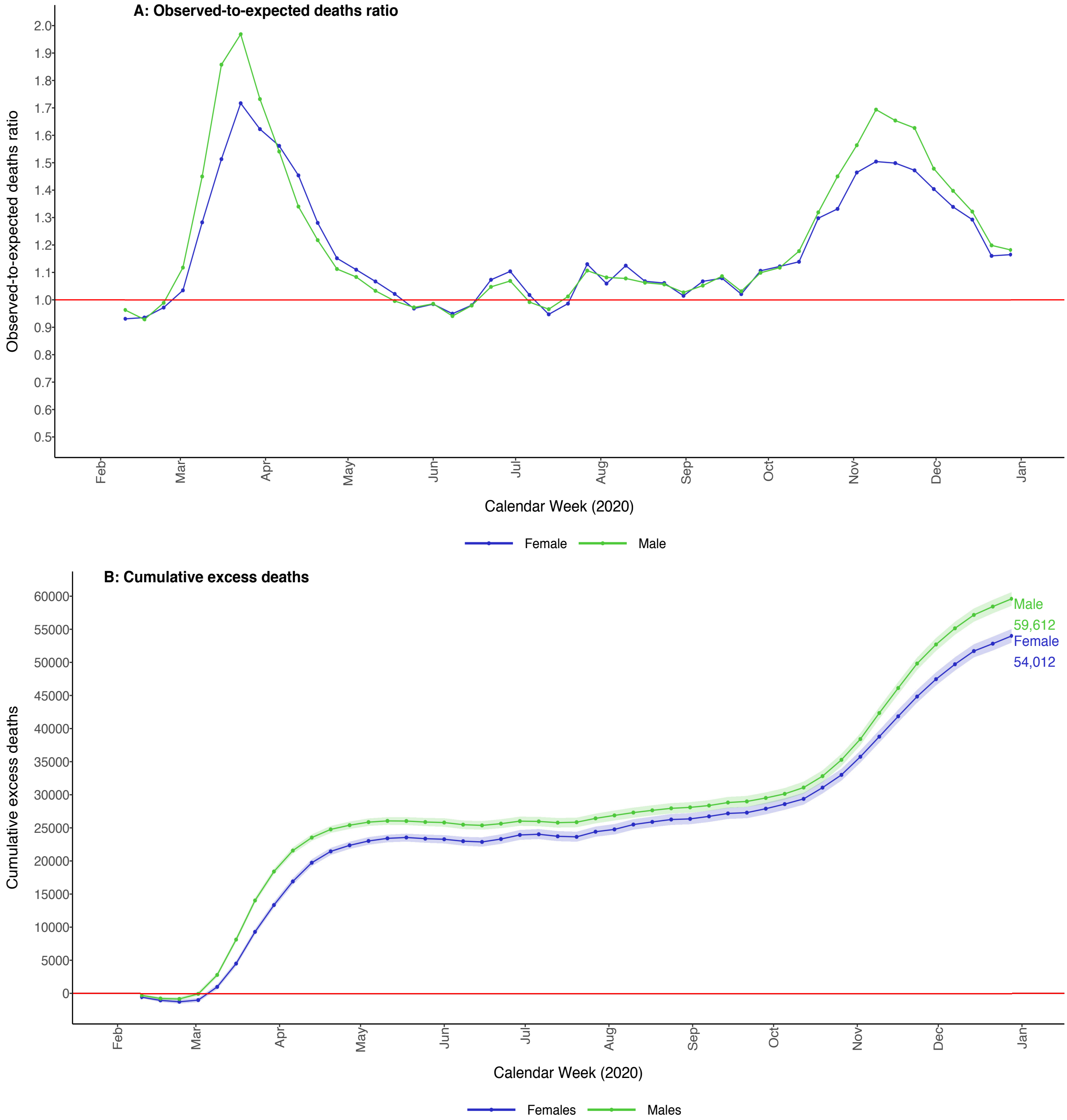
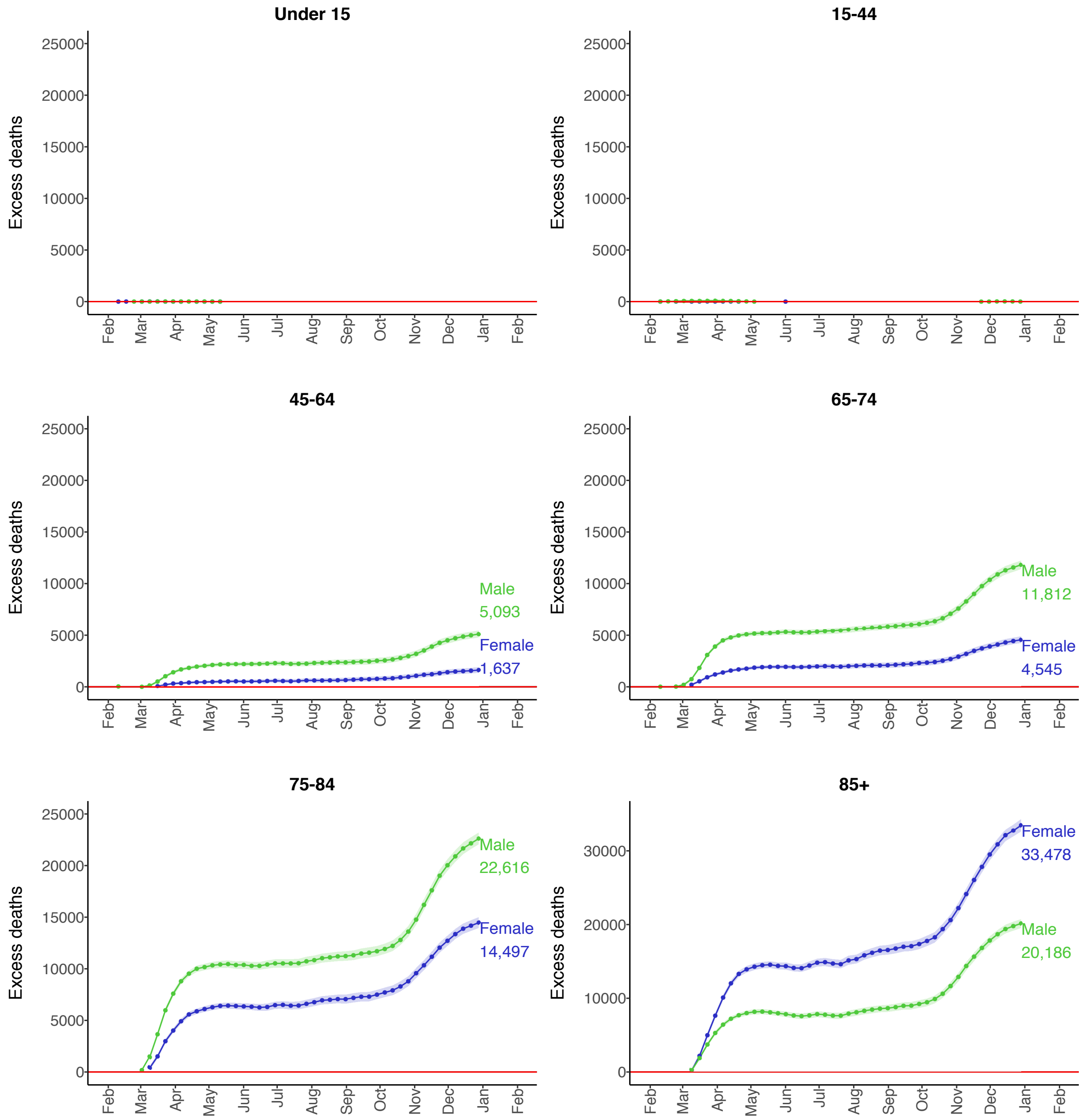


Figure 4.13: Observed-to-expected deaths ratio and cumulative excess weekly deaths Italy - 2020



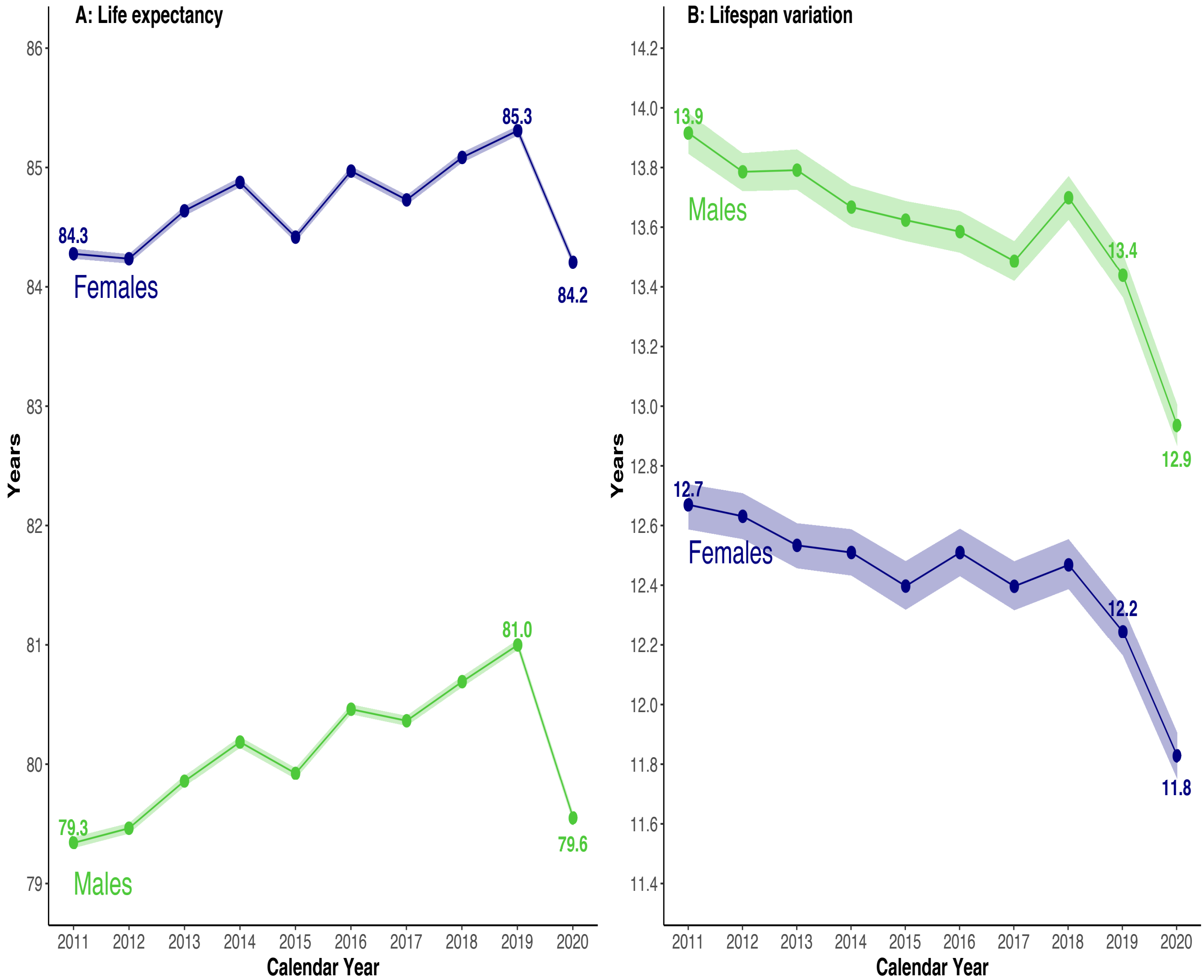
Notes: Shaded areas in Panel B represent 95% prediction intervals, constructed by sampling weekly deaths count from Poisson distribution. Excess deaths are defined as the difference between total observed deaths and expected (baseline) deaths

Figure 4.14: Cumulative excess weekly deaths by sex and age
Italy - 2020



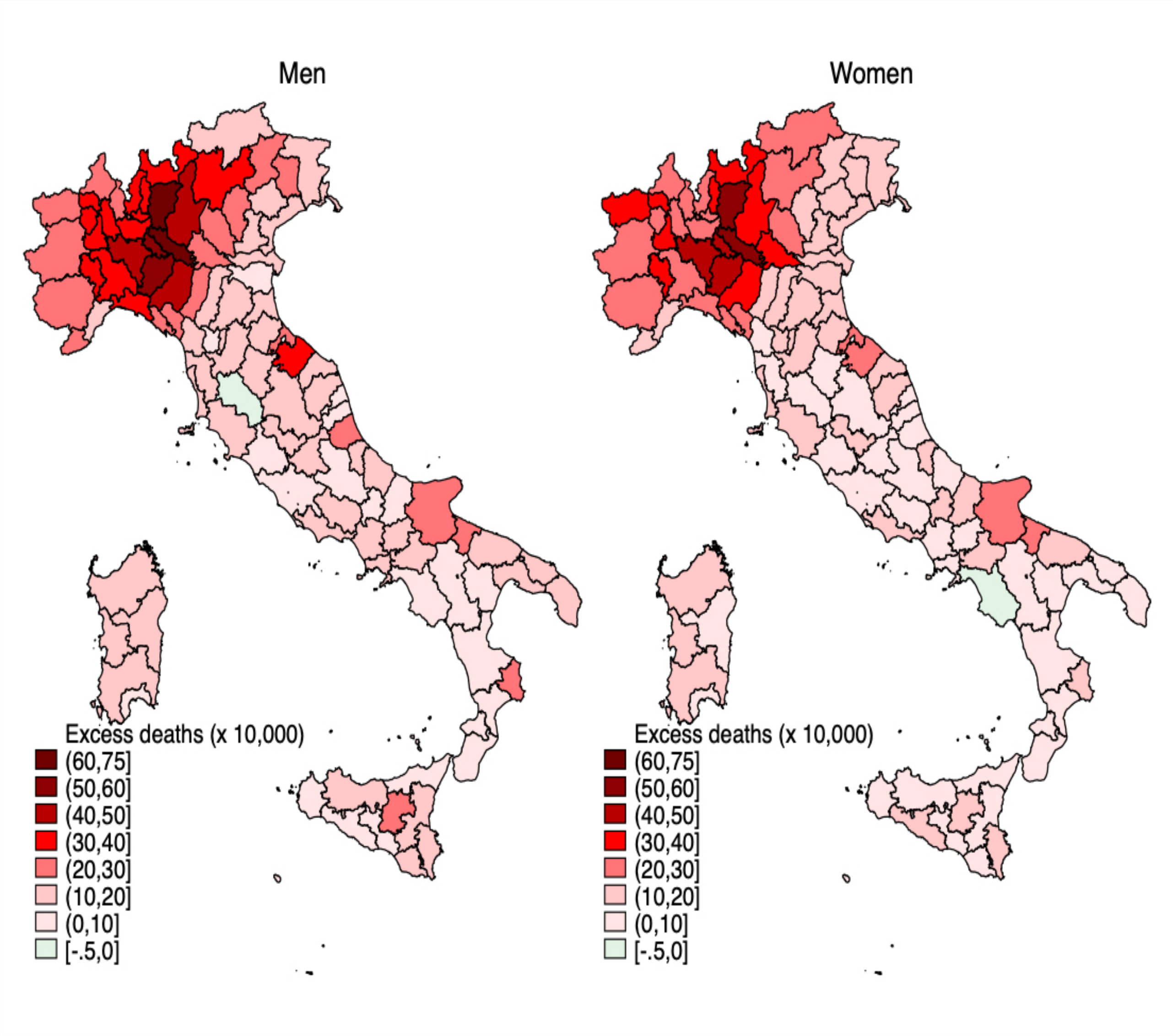
Notes: Shaded areas represent 95% prediction intervals, constructed by sampling weekly deaths count from Poisson distribution. Excess deaths are defined as the difference between total observed deaths and expected (baseline) deaths

Figure 4.15: Life expectancy and lifespan variation at birth
Italy - 2011-2020



Notes: Shaded areas represent 95% confidence intervals, constructed by assuming yearly deaths follow a binomial distribution

Figure 4.16: Excess mortality rate (per 10,000)
Italian provinces - 2020



Notes: Excess mortality is calculated as the weighted sum of age group-specific excess mortality rate. Excess mortality rate for each age group is calculated as the ratio between excess deaths (difference between observed and expected deaths) divided by population exposure. Sex- and age-group specific weights were obtained using the 2020 Italian population as standard.

Figure 4.17: Life expectancy and lifespan variation at birth
Italian provinces - 2020 vs 2017-2019 average

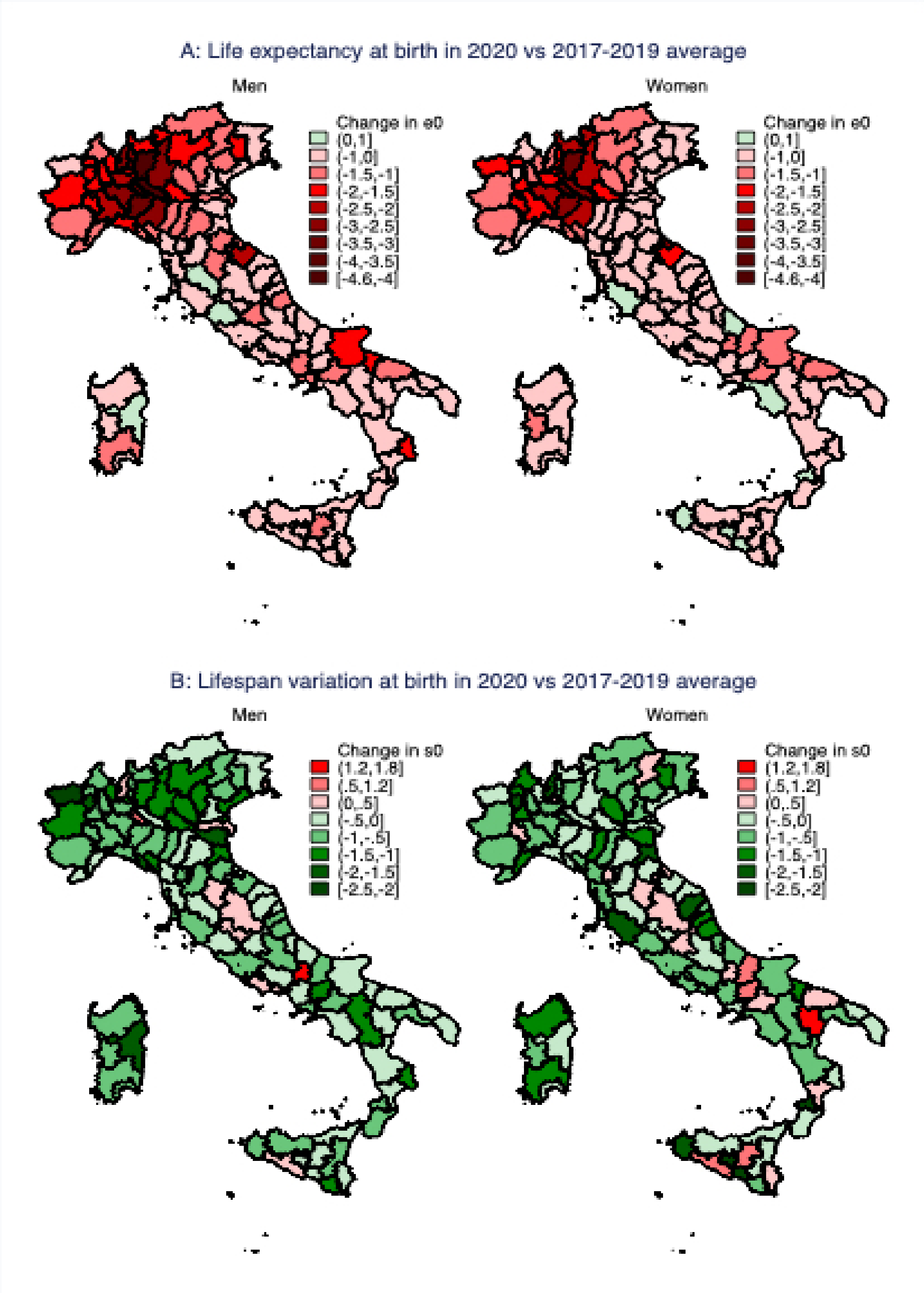


Figure 4.18: Impact on lifespan variation at birth vs impact on life expectancy at birth in Italian provinces 2020

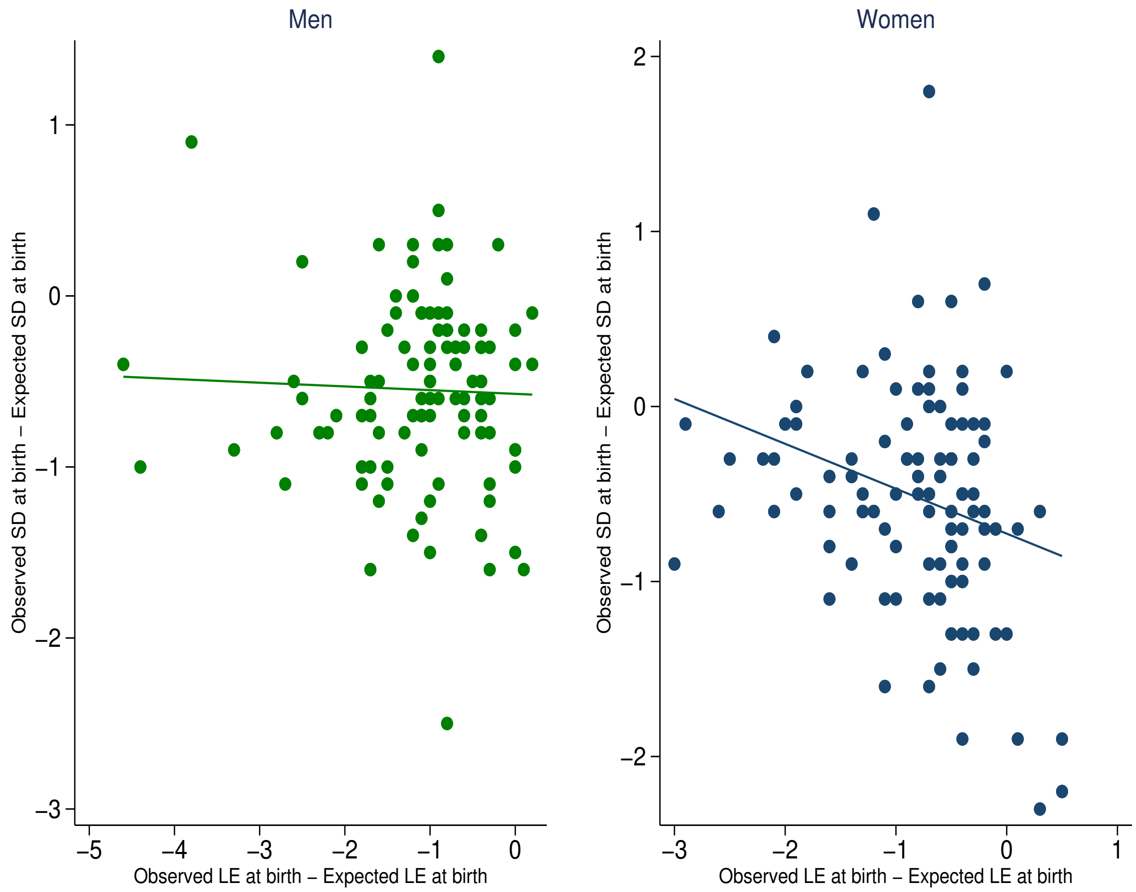
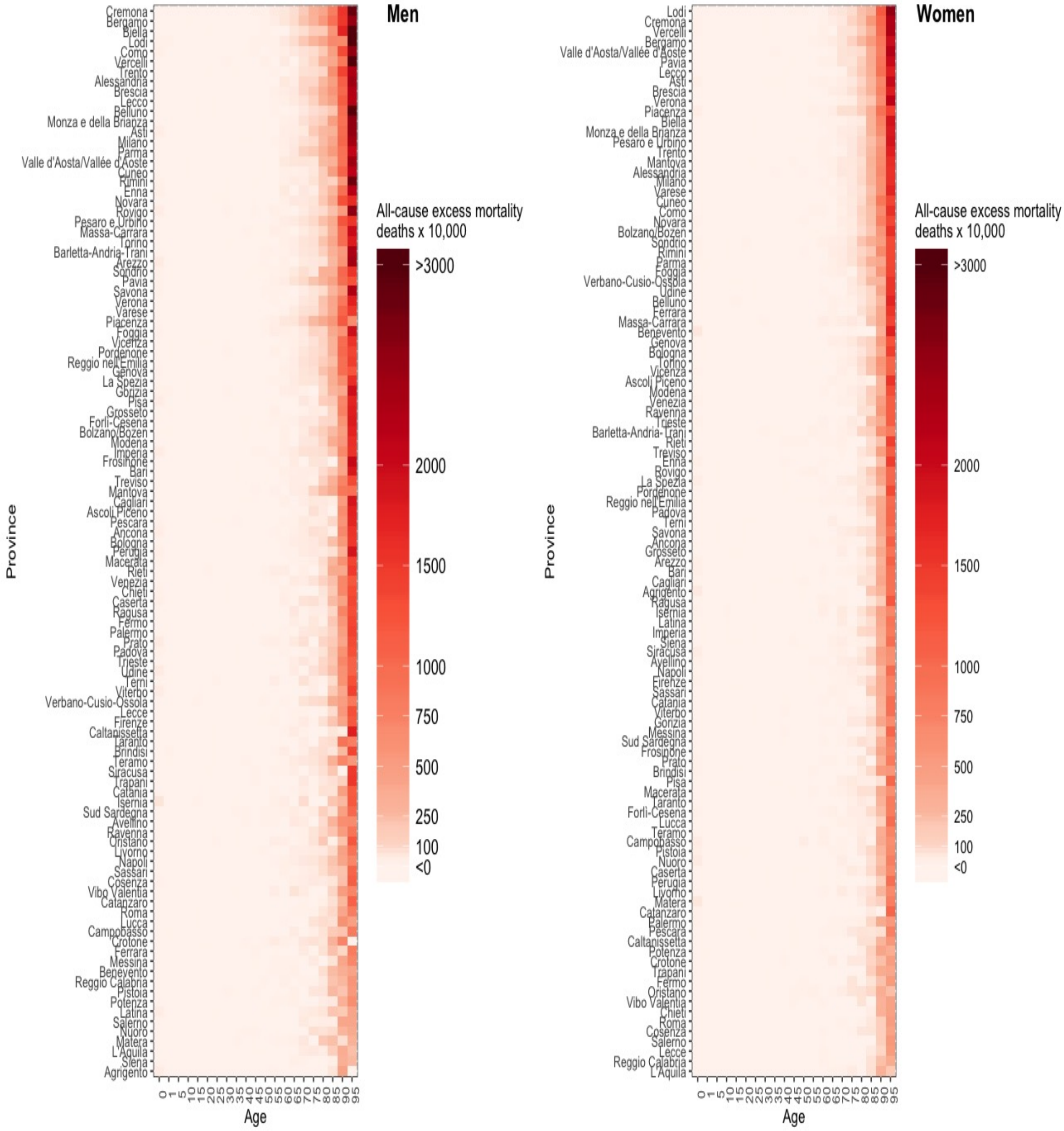
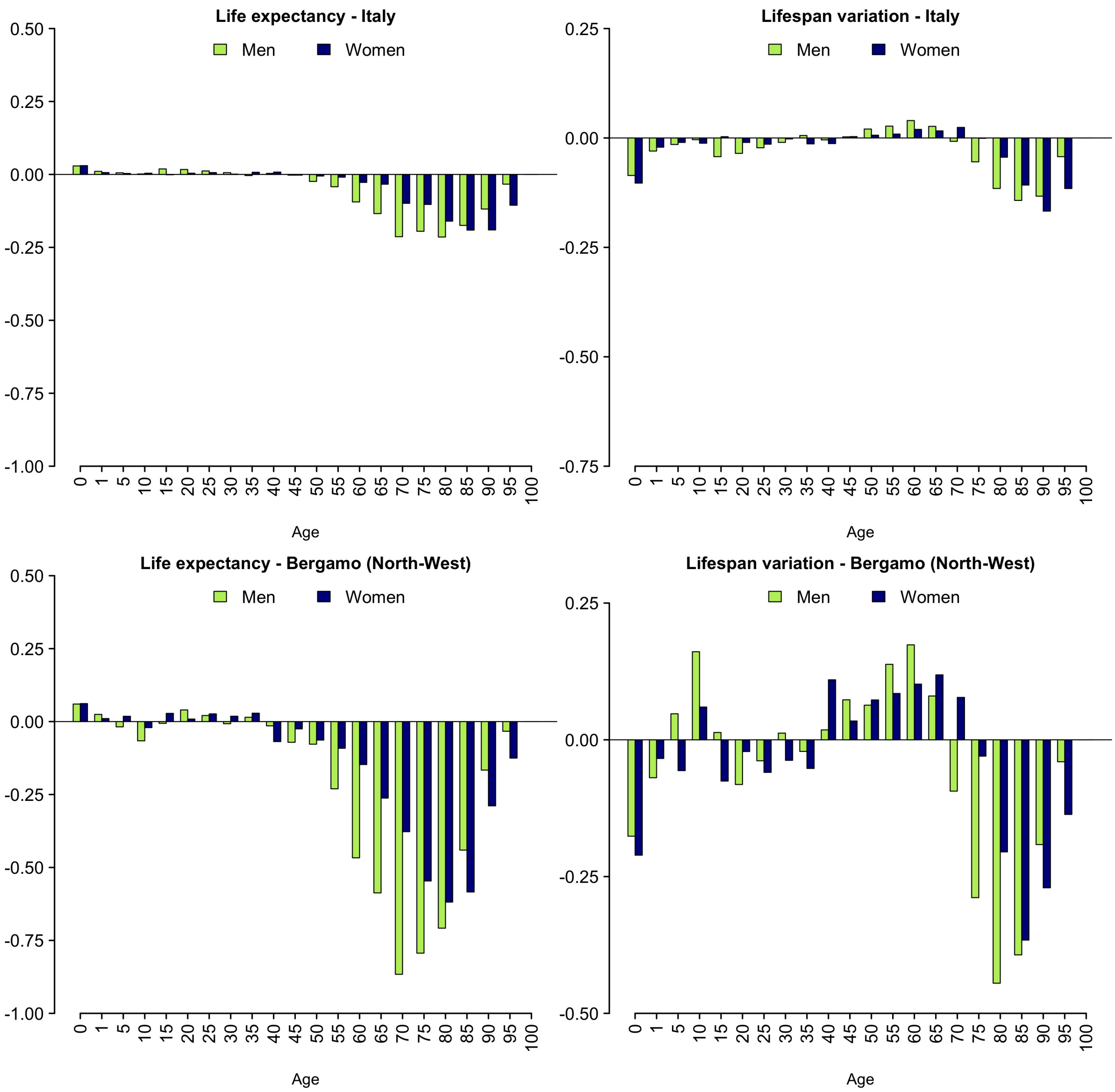


Figure 4.19: Excess mortality rate per 10,000 by age
Italian provinces, 2020 vs 2017-2019 average



Notes. Excess mortality rate calculated as the difference between observed and expected mortality rate. Provinces are ordered by (total) age-standardized excess mortality rate.

Figure 4.20: Age-specific contribution to difference between observed life expectancy and lifespan variation at birth in 2020 vs 2017-2019 average



Tables Chapter 4

Table 4.1: Excess deaths, life expectancy and lifespan inequality (observed vs baseline) in 1957-1958 and 1969-19670, Italy

Excess deaths (per 10,000)		
Year	Men	Women
1957	9.4846 [9.4775-9.5284]	7.5877 [7.5519-7.6236]
1958	2.2636 [2.2241-2.2819]	1.6859 [1.6724-1.7212]
1969	6.7434 [6.7228-6.7808]	3.7502 [3.7068-3.7704]
1970	0.8912 [0.8921-0.9027]	0.4310 [0.4091-0.4546]

Life expectancy at birth				
	Men		Women	
	Observed	Baseline	Observed	Baseline
1957	65.63 [65.56- 65.71]	66.91 [66.84- 66.97]	70.18 [70.11- 70.25]	71.28 [71.21- 71.34]
1958	66.64 [66.56- 66.71]	67.01 [66.94- 67.08]	71.23 [71.16- 71.29]	71.50 [71.43- 71.57]
1969	67.93 [67.87- 67.99]	68.73 [68.67- 68.78]	73.84 [73.78- 73.91]	74.40 [74.35- 74.46]
1970	68.78 [68.72- 68.84]	68.82 [68.76- 68.88]	74.54 [74.49- 74.60]	74.59 [74.53- 74.65]

Lifespan variation at birth (standard deviation in age-at-death)				
	Men		Women	
	Observed	Baseline	Observed	Baseline
1957	22.48 [22.42- 22.56]	22.46 [22.38- 22.53]	21.74 [21.67- 21.82]	21.61 [21.54- 21.68]
1958	22.41 [22.38- 22.53]	22.26 [21.67- 21.82]	21.48 [21.54- 21.68]	21.41 [21.33- 21.49]
1969	19.72 [19.65- 19.78]	19.82 [19.76- 19.89]	18.64 [18.57 - 18.72]	18.56 [18.49- 18.65]
1970	19.55 [19.47- 19.61]	19.65 [19.58- 19.73]	18.33 [18.25 - 18.41]	18.37 [18.29- 18.45]

Notes. Excess mortality is calculated as the difference between observed and baseline age-standardized mortality rates, per 10,000.

Table 4.2: Predictors of excess mortality in 1957 and 1969 in Italian provinces

Excess mortality	(1)	(2)	(3)
Population density	0.00530**	0.0106***	0.000271
	(0.00227)	(0.00350)	(0.00272)
Share of LF in industry	0.239	-0.275	1.201**
	(0.414)	(0.470)	(0.597)
Women	-8.020***	-4.221	-11.78**
	(2.822)	(3.864)	(4.485)
Constant	25.01	29.61	-6.233
	(21.48)	(25.86)	(29.10)
Observations	366	182	184
R-squared	0.258	0.483	0.544
Years	1957-1959	1957	1969
Region FE	Yes	Yes	Yes
Year FE	Yes	No	No

Notes. Results from OLS regression. Dependent variable: excess deaths/100,000 population by sex, all age groups combined, in 1957 and 1969 in Italian provinces. In each year, excess deaths are age-standardized using the Italian population as standard. Province-level predictors include (i) population density (number of inhabitants per km2), and (ii) employment share in industry (% of total employment). Both predictors refer to census years 1961 (for pandemic year 1957) and 1971 (for pandemic year 1969) and are publicly available on ISTAT website. Robust standard errors are reported in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Bibliography

Aassve, A., Alfani, G., Gandolfi, F., & Le Moglie, M. (2021). Epidemics and Trust: The Case of the Spanish Flu. *Health Economics*, *30*(4), 840-857. doi: https://doi.org/10.1002/hec.4218

Abbott, A. (1995). Sequence Analysis: New Methods for Old Ideas. *American Journal of Sociology*, *21*(1), 93-113.

Aburto, J.-M., Kashyap, R., Schöley, J., Angus, C., Ermisch, J., Mills, M. C., & Dowd, J. B. (2021). Estimating the burden of the COVID-19 pandemic on mortality, life expectancy and lifespan inequality in England and Wales: a population-level analysis. *Journal of Epidemiology and Community Health*, *0*, 1-6. doi: 10.1136/jech-2020-215505

Aburto, J.-M., Schöley, J., Kashnitsky, I., Zhang, L., Rahal, C., Missov, T. I., ... Kashyap, R. (2021). Quantifying impacts of the COVID-19 pandemic through life expectancy losses: a population-level study of 29 countries. *medRxiv*. Retrieved from `https://www.medrxiv.org/content/early/2021/04/06/2021.03.02.21252772` doi: 10.1101/2021.03.02.21252772

Aburto, J.-M., Villavicencio, F., Basellini, U., Kjærgaard, S., & Vaupel, J. W. (2020). Dynamics of life expectancy and life span equality. *Proceedings of the National Academy of Sciences*, *117*(10), 5250-5259. doi: https://doi.org/10.1073/pnas.1915884117

Acosta, E., Hallman, S. A., Dillon, L. Y., Ouellette, N., Bourbeau, R., Herring, D. A.,

. . . Gagnon, A. (2019). Determinants of influenza mortality trends: Age-period-cohort analysis of influenza mortality in the united states, 1959–2016. *Demography*, *56*, 1723–1746.

Ahlbom, A. (1993). *Biostatistics for Epidemiologists.* Lewis Publishers.

Aimone Gigio, L., Camussi, S., & Maccarrone, V. (2021). *Changes in the employment structure and in job quality in Italy: a national and regional analysis.* Retrieved from https://www.bancaditalia.it/pubblicazioni/qef/2021-0603/QEF_603_21.pdf?language_id=1 (Questioni di Economia e Finanza - Bank of Italy Occasional Papers, N. 603)

Aisenmbrey, S., & Fasang, A. (2010). New Life for Old Ideas: The "Second Wave" of Sequence Analysis. Bringing the "Course" Back Into the Life Course. *Sociological Methods & Research*, *38*(3), 420-462.

Almond, D. (2006). Is the 1918 influenza pandemic over? Long-term effects of in utero influenza exposure in the post-1940 U.S. population. *Journal of Political Economy*, *114*(4), 672-711.

Alvarez, J.-A., Kallestrup-Lamb, M., Kjærgaard, S., & Vaupel, J. W. (2020). *Linking retirement age to life expectancy does not lessen the demographic implications of unequal lifespans.* (Working Paper)

Andreasen, V., Viboud, C., & Simonsen, L. (2008). Epidemiologic characterization of the 1918 influenza pandemic summer wave in Copenhagen: implications for pandemic control strategies. *Journal of Infectious Diseases*, *197*(2), 270-278. doi: 10.1086/524065

Andreev, E., & Shkolnikov, V. (2010). *Spreadsheet for calculation of confidence limits for any life table or healthy-life table quantity.* (Rostock: Max Planck Institute for

Demographic Research. https://www.demogr.mpg.de/en/publications_databases_-6118/publications_1904/mpidr_technical_reports/spreadsheet_for_calculation_-of_confidence_limits_for_any_life_table_or_healthy_life_table_quantity_3853 (accessed April 28, 2020).)

Andreev, E., & Shkolnikov, V. (2012). *An Excel spreadsheet for the decomposition of a difference between two values of an aggregate demographic measure by stepwise replacement running from young to old ages, note =.*

Ardito, C., Costa, G., & Leombruni, R. (2019). Differenze sociali nella salute ed equità del sistema pensionistico italiano / Socioeconomic differences in health and equity of the Italian pension system. *La Rivista delle Politiche Sociali / The Italian Journal of Social policies*, *3*, 13-25.

Ardito, C., D'Errico, A., Leombruni, R., Zengarini, N., Ricceri, F., & Costa, G. (2021). *The evolution of longevity inequality in Italy and concerns for the pension system.* (Working paper - VISITINPS Project)

Auerbach, A. J., Charles, K. K., Coile, C. C., Gale, W., Goldman, D., Lee, R., ... Wongr, R. (2017). How the Growing Gap in Life Expectancy May Affect Retirement Benefits and Reforms. *The Geneva Papers*, *42*, 475-499.

Ayuso, M., Bravo, J. M., & Holzmann, R. (2017). On the Heterogeneity in Longevity among Socioeconomic Groups: Scope, Trends, and Implications for Earnings-Related Pension Schemes. *Global Journal of Human Social Sciences-Economics*, *17*(1), 33-58.

Ayuso, M., Bravo, J. M., & Holzmann, R. (2020). Getting life expectancy estimates right for pension policy: period versus cohort approach. *Journal of Pension Economics and Finance*, 1-20. doi: 10.1017/S1474747220000050

Barban, N., & Billari, F. (2012). Classifying life course trajectories: a comparison of latent class and sequence analysis. *Journal of the Royal Statistical Society*, *61*(5), 765-784.

Barban, N., de Luna, X., Lundholm, E., Svensson, I., & Billari, F. C. (2017). Causal Effects of the Timing of Life-Course Events: Age at Retirement and Subsequent Health. *Sociological Methods and Research*, *XX*(X), 1-34.

Belloni, M., Alessie, R., Kalwij, A., & Marinacci, C. (2012). Lifetime Income and Old Age Mortality Risk in Italy over Two Decades. *Demographic Research*, *29*(45), 1261-1298. doi: https://doi.org/10.4054/DemRes.2013.29.45

Belloni, M., & Maccheroni, C. (2013). Actuarial Fairness When Longevity Increases: An Evaluation of the Italian Pension System. *The Geneva Papers*, *38*, 638-674.

Benach, J., & Muntaner, C. (2007). Precarious Employment and Health: Developing a Research Agenda. *Journal of Epidemiology and Community Health*, *61*(4), 276–277.

Benach, J., Vives, A., Amable, M., Vanroelen, C., Tarafa, G., & Muntaner, C. (2015). Precarious Employment: Understanding an Emerging Social Determinant of Health. *Annual Review of Public Health*, *35*, 229–253. doi: 10.1146/annurev-publhealth-032013-182500

Bertuccio, P., Alicandro, G., Sebastiani, G., Zengarini, N., Costa, G., La Vecchia, C., & Frova, L. (2018). Mortality by occupation-based social class in Italy from 2012 to 2014. *International Journal of Public Health*, *63*, 865–874. doi: 10.1007/s00038-018-1149-8

Bessudnov, A., McKee, M., & Stuckler, D. (2011). Inequalities in male mortality by occupational class, perceived status and education in Russia, 1994–2006. *European Journal of Public Health*, *22*(3), 332–337. doi: https://doi.org/10.1093/eurpub/ckr130

Blakely, T. A., Collings, S. C. D., & Atkinson, J. (2003). Unemployment and suicide. Evidence for a causal association? *Journal of Epidemiology Community Health*, *57*, 5594-600.

Blangiardo, G. (2020). *Scenari sugli effetti demografici di Covid-19 per l'anno 2020.* Retrieved from `https://www.istat.it/it/files/2020/04/Report-Presidente_Finale.pdf`

Blangiardo, N., Cameletti, M., Pirani, N., Corsetti, G., Battaglini, M., & Baio, G. (2020). Estimating weekly excess mortality at sub-national level in Italy during the COVID-19 pandemic. *PLOS One*, *15*(10), e0240286. doi: 10.1371/journal.pone.0240286

Bloemen, H., Hochguertel, S., & Zweerink, J. (2019). The Effect of Incentive-Induced Retirement on Spousal Retirement Rates: Evidence from a Natural Experiment. *Economic Inquiry*, *57*(2), 910-930.

Bloise, F., & Raitano, M. (2021). Intergenerational earnings persistence in italybetween actual father–son pairs accounting forlifecycle and attenuation bias. *Oxford Bulletin of Economics and Statistics*, *83*(1), 88-114.

Bosworth, B., Burtless, G., & Zhang, K. (2016). *Later Retirement, Inequality in Old Age, and the in Old Age, and the Growing Gap in Longevity Between Rich and Poor.* Economic Studies at Brookings.

Bourgignon, F. (2018). World Changes in Inequality: an Overview of Facts, Causes, Consequences, and Policies. *CESifo Economic Studies*, *64*(3), 345-370. doi: https://doi.org/10.1093/cesifo/ifx028

Brass, W. (1971). On the scale of mortality. In W. Brass (Ed.), *Biological aspect of demography* (p. 69-110). London: Taylor and Francis.

Brugiavini, A., & Peracchi, F. (2003). Social security wealth and retirement decisions in Italy. *Labour*, *17*(s1), 79-114. doi: https://doi.org/10.1111/1467-9914.17.specialissue.4

Brønnum-Hansen, H. (2017). Socially disparate trends in lifespan variation: a trend study on income and mortality based on nationwide Danish register data. *BMJ Open*, *7*. doi: https://doi.org/10.1136/bmjopen-2016-014489

Brønnum-Hansen, H., & Baadsgaard, M. (2012). Widening social inequality in life expectancy in Denmark. A register-based study on social composition and mortality trends for the Danish population. *BMC Public Health*, *12*. doi: https://doi.org/10.1186/1471-2458-12-994

Burdorf, A. (2010). Is Early Retirement Good for Your Health? *BMJ (Clinical Research Ed.)*, *341*, c6089.

Cambois, E., Brønnum-Hansen, H., Hayward, M., & Nusselder, W. (2020). Monitoring social differentials in health expectancies. In C. Jagger, E. Crimmins, Y. Saito, Y. R. D. Carvalho, H. V. Oyen, & J. Robine (Eds.), *International Handbook of Health Expectancies. International Handbooks of Population, vol. 9* (p. 45-66). Springer.

Caroli, E., & Godard, M. (2014). Does job insecurity deteriorate health? *Health Economics*, *25*(2), 131-147.

Carta, F., & Dephilippis, M. (2021). *Working horizon and labour supply: the effect of raising the full retirement age on middle-aged individuals.* Retrieved from `https://www.bancaditalia.it/pubblicazioni/temi-discussione/2021/2021-1314/en_Tema_1314.pdf` (Bank of Italy Working Papers, N. 1314)

Case, A., & Deaton, A. (2021). Life expectancy in adulthood is falling for those without a ba degree, but as educational gaps have widened, racial gaps have narrowed. *Proceedings of the National Academy of Sciences*, *118*(11). doi: 10.1073/pnas.2024777118

Caselli, G., & Lipsi, R. M. (2018). Survival inequalities and redistribution in the Italian pension system. *Vienna Yearbook of Population Research*, *16*, 1-28. doi: https://

doi.org/10.1553/populationyearbook2018s083

Checchi, D., & Leonardi, M. (2015). *abour market measures in Italy 2008–13 : the crisis and beyond.* Retrieved from http://www.ilo.int/wcmsp5/groups/public/---dgreports/---inst/documents/publication/wcms_449929.pdf (International Labour Office, Research Department. – Geneva: ILO)

Chetty, R., Stepner, M., Abraham, S., Lin, S., Scuderi, B., Turner, N., ... Cutler, D. (2016). Association Between Income and Life Expectancy in the United States. *Journal of American Medical Association*, *315*(16), 1750-1766. doi: https://doi.org/10.1001/jama.2016.4226

Cheung, S. L. K., Robine, J.-M., Tu, E. J.-C., & Caselli, G. (2005). Three dimensions of the survival curve: Horizontalization, verticalization, and longevity extension. *Demography*(42), 243-258. doi: https://doi.org/10.1353/dem.2005.0012

Chiang, C. (1984). *The Life Table and Its Applications, Original edition.* Malabar, FL: Krieger Publishing Company.

Chowell, G., Erkoreka, A., Viboud, C., & Echeverri-Dávila, B. (2014). Spatial-temporal excess mortality patterns of the 1918–1919 influenza pandemic in Spain. *BMC Infectious Diseases*, *14*, 371.

Clark, A. E. (2003). Unemployment as a Social Norm: Psychological Evidence from Panel Data. *Journal of Labor Economics*, *21*(2), 323-351.

Clark, A. E., Georgellis, Y., & Sanfey, P. (2001). Scarring: The Psychological Impact of Past Unemployment. *Economica*, *68*, 221-241.

Clark, A. E., & Lepinteur, A. (2019). The causes and consequences of early-adult unemployment: Evidence from cohort data.

Contini, B., & Leombruni, R. (2006). From work to retirement: a tale of bumpy routes. *Review of Political Economy*, *18*(3), 359-378.

Costa, G., Faggiano, F., Cadum, E., Lagorio, S., Arcà, M., Farchi, G., ... Pagnanelli, F. (1994). Le differenze sociali nella mortalità in Italia. In G. Costa & F. Faggiano (Eds.), *L'equità nella salute in Italia.* Milano: Franco Angeli.

Costa, G., Stroscia, M., Zengarini, N., & Demaria, M. (2017). *40 anni di salute a Torino. Spunti per leggere i bisogni e i risultati delle politiche.* Milano: Inferenze.

Cox, D. (1972). Regression Models and Life-Tables. *Journal of the Royal Statistical Society. Series B (Methodological)*, *34*(2), 187-220.

Cristia, J. P. (2009). Rising mortality and life expectancy differentials by lifetime earnings in the United States. *Journal of Health Economics*, *28*, 984-995. doi: https://doi.org/10.1016/j.jhealeco.2009.06.003

Cutler, D., Deaton, A., & LLeras-Muney, A. (2006). The Determinants of Mortality. *Journal of Economic Perspectives*, *20*(2), 97-120. doi: https://doi.org/10.1257/jep.20.3.97

Demetrius, L. (1978). Adaptive value, entropy and survivorship curves. *Nature*, *275*, 213-214. doi: https://doi.org/10.1038/275213a0

Devillanova, C., Raitano, M., & Struffolino, E. (2019, 5). Longitudinal employment trajectories and health in middle life: Insights from linked administrative and survey data. *Demographic Research*, *40*(5), 1375–1412.

Di Novi, C., & Marenzi, A. (2019). The smoking epidemic across generations, genders, and educational groups: A matter of diffusion of innovations. *Economics & Human Biology*, *33*, 155-168. doi: https://doi.org/10.1016/j.ehb.2019.01.003

Dooley, D., Fielding, J., & Levi, L. (1996). Health and Unemployment. *Annual Review of Public Health*, *17*, 449-465.

Dougherty, R., Edelman, A., & Hyman, J. (1989). Nonnegativity-, monotonicity-, or convexity-preserving cubic and quintic hermite interpolation. *Mathematics of Computation*, *52*(186), 471-493.

Efron, B. (1977). The efficiency of Cox's likelihood function for censored data. *Journal of the American Statistical Association*, *72*(359), 557-565.

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

Eurofound. (2016). *Extending working lives through flexible retirement schemes: Partial retirement.* Luxembourg: Publications Office of the European Union.

Fadlon, I., & Nielsen, T. H. (2020). Family Labor Supply Responses To Severe Health Shocks. , *Forthcoming*.

Fasang, A. E. (2010). Retirement: Institutional Pathways and Individual Trajectories in Britain and Germany. *Sociological Research Online*, *15*(2).

Fasang, A. E. (2012, 3). Retirement Patterns and Income Inequality. *Social Forces*, *90*(3), 685-711.

Felice, E. (2018). The Socio-Institutional Divide: Explaining Italy's Long-Term Regional Differences. *Journal of Interdisciplinary History*, *XLIX*(1), 43-70. doi: https://doi.org/10.1162/jinh_a_01231

Gabadinho, A., Ritschard, G., Müller, N., & Studer, M. (2011). Analyzing and Visualizing State Sequences in R with TraMineR. *Journal of Statistical Software*, *40*(4), 1-37.

Galasso, V., Pons, V., Profeta, P., Becher, M., Brouard, S., & Foucault, M. (2020). Gender differences in COVID-19 attitudes and behavior: Panel evidence from eight countries. *Proceedings of the National Academy of Sciences*, *117*(44), 27285-27291. doi: 10.1073/pnas.2012520117

Galizzi, M., & Ghislandi, S. (2020). *Bergamo's response to the coronavirus pandemic.* (Electronic Resource: https://www.cambridge.org/core/blog/.)

Gatto, M., Bertuzzo, E., Mari, L., Miccoli, S., Carraro, L., Casagrandi, R., & Rinaldo, A. (2020). Spread and dynamics of the COVID-19 epidemic in Italy: Effects of emergency containment measures. *Proceedings of the National Academy of Sciences*, *117*(19). doi: 10.1073/pnas.2004978117.

Gauthier, J.-A., Wider, E., Bucher, P., & Notredame, C. (2009). How much does it cost? Optimization of costs in sequence analysis of social science data. *Sociological Methods & Research*, *38*(1), 197-231.

Gavrilov, L. A., & Gavrilova, N. S. (2011). Mortality measurement at advanced ages: A study of the Social Security Administration Death Master File. *North American Actuarial Journal*, *15*(3), 432-447.

Geyer, S. (2006). Education, income, and occupational class cannot be used interchangeably in social epidemiology. Empirical evidence against a common practice. *Journal of Epidemiology and Community Health*, *60*, 804-810. doi: 10.1136/jech.2005.041319

Ghislandi, S., Sanderson, W., & Scherbov, S. (2019). A Simple Measure of Human Development: The Human Life Indicator. *Population and Development Review*, *45*, 219-233.

Giupponi, G. (2019). *When Income Effects are Large: Labor Supply Responses and the Value of Welfare Transfers.* (CEP Discussion Papers, dp1651, Centre for Economic

Performance, LSE)

Goldman, N., & Lord, G. (1986). A new look at entropy and the life table. *Demography*, *23*, 275-282. doi: https://doi.org/10.2307/2061621

Gompertz, B. (1825). On the nature of the function expressive of the law of human mortality, and on a new mode of determining the value of life contingencies. *Philosophical Transactions of the Royal Society of London*, *115*, 513-585.

Guillot, M., & Payne, C. F. (2019). Tracking progress in mean longevity: The Lagged Cohort Life Expectancy (LCLE) approach. *Population Studies*, *73*(3), 405-421. doi: 10.1080/00324728.2019.1618480

Haan, P., Kemptner, D., & Luther, H. (2020). The Rising Longevity Gap by Lifetime Earnings: Distributional Implications for the Pension System. *The Journal of the Economics of Ageing*, *17*(100199). doi: https://doi.org/10.1016/j.jeoa.2019.100199

Haider, S., & Solon, G. (2006). Life-Cycle Variation in the Association between Current and Lifetime Earnings, volume = 96, number = 4, journal = American Economic Review, doi = 10.1257/aer.96.4.1308. , 1038-1320.

Hayward, M. D., & Grady, W. R. (1990). Work and Retirement among a Cohort of Older Men in the United States, 1966–1983. *Demography*, *27*(3), 337-356.

Henderson, D. A., Inglesby, T. V., Toner, E., & Nuzzo, J. B. (1969). Origin and progress of the 1968-69 Hong Kong influenza epidemic. *Bulletin World Health Organization*, *41*(3-4-5), 343–348.

Hult, C., Stattin, M., Janlert, U., & Jarvholm, B. (2010). Timing of Retirement and Mortality — A Cohort Study of Swedish Construction Workers. *Social Science Medicine*, *79*, 1480-1486.

Italian National Institute of Statistics. (1985). *Nota medological a Popolazione e bilanci demografici per sesso, età e regione. Ricostruzione per gli anni 1972-1981.* (Supplemento al Bollettino Mensile di Statistica, n. 14)

Italian National Institute of Statistics. (2020). *Conti Economici Territoriali | Anni 2017-2019.* Retrieved from `https://www.istat.it/it/files//2020/12/REPORT-CONTI-TERRITORIALI_2019.pdf`

Italian National Institute of Statistics. (2021). *Indicatori Demografici | Anno 2020.* Retrieved from `https://www.istat.it/it/files//2021/05/REPORT_INDICATORI-DEMOGRAFICI-2020.pdf`

Jenkins, S. (2005). *Survival analysis.* (Unpublished Manuscript)

Jin, R. L., Shah, C. P., & Svoboda, T. J. (1995). The impact of unemployment on health: a review of the evidence. *Canadian Medical Association Journal*, *153*(5), 529-540.

Johnson, N., & Mueller, J. (2002). Updating the accounts: global mortality of the 1918-1920 "Spaniffsh" influenza pandemic. *Bulletin Historical Medicine*, *76*(1), 105-115.

Johnson, N. J., Sorlie, P. D., & Backlund, E. (1999). The impact of specific occupation on mortality in the U.S. National Longitudinal Mortality Study. *Demography*, *36*, 355–367.

Katikireddi, S. V., Leyland, A. H., McKee, M., Ralston, K., & Stuckler, D. (2017). Patterns of mortality by occupation in the UK, 1991–2011: a comparative analysis of linked census and mortality records. *The Lancet*, *2*(11), E501-E512. doi: https://doi.org/10.1016/S2468-2667(17)30193-7

Kaufman, L., & Rousseeuw, P. J. (2005). *Finding groups in data: An introduction to cluster analysis.* Hoboken, NJ, USA: John Wiley Sons.

Keyfitz, N. (1977). *Introduction to the Mathematics of Population: With Revisions.* Addison-Wesley Series in Behavioral Science: Quantitative Methods, Addison-Wesley.

Keyfitz, N., & Golini, A. (1975). Mortality comparisons: The male-female ratio. *Genus*, *31*, 1-34.

Kiebele, E., Jasilionis, D., & Shkolnikov, V. (2013). Widening socioeconomic differences in mortality among men aged 65 years and older in Germany. *Journal of Epidemiology Community and Health*, *67*, 453-457.

Kopczuk, W., Saez, E., & Song, J. (2010). Earnings inequality and mobility in the United States: Evidence from Social Security Data since 1937. *Quarterly Journal of Economics*, *125*(1), 91-128.

Krieger, N., Chen, J., & Waterman, P. (2020). Excess mortality in men and women in Massachusetts during the COVID-19 pandemic. *The Lancet*, *395*, 1829.

Lallo, C., & Raitano, M. (2018). Life Expectancy Inequalities in the Elderly by Socioeconomic Status: Evidence from Italy. *Population Health Metrics*, *16*(7). doi: https://doi.org/10.1186/s12963-018-0163-7

Leombruni, R., D'Errico, A., Stroscia, M., Zengarini, N., & Costa, G. (2015). Inequalities among pre-retirement individuals: life expectancy variations and implications for social security policies. *POLITICHE SOCIALI / Social policies*, *3*, 461-479. doi: https://doi.org/10.7389/81766

Leombruni, R., Richiardi, M., Demaria, M., & Costa, G. (2010). Life expectancy, strenuous work and pension system's fairness. First evidence from the Work Histories Italian Panel. *Epidemiologia e Prevenzione*, *4*, 150-158.

Leser, C. (1955). Variations in mortality and life expectation. *Population Studies*, *9*, 67-71. doi: https://doi.org/10.2307/2172342

Lesnard, L. (2006). Setting Cost in Optimal Matching to Uncover Contemporaneous SocioTemporal Patterns. *Sociological Methods Research*, *38*(3), 389–419.

Lewer, D., Jayatunga, W., Aldridge, R. W., Edge, C., Marmot, M., Story, A., & Hayward, P. A. (2020). Premature mortality attributable to socioeconomic inequality in England between 2003 and 2018: an observational study, volume = 5, number = 1, journal = The Lancet Public Health, doi = . , E33-E41.

Lorenti, A., Dudel, C., & Myrskylä, M. (2019). The Legacy of the Great Recession in Italy: A Wider Geographical, Gender, and Generational Gap in Working Life Expectancy. *Social Indicators Research volume*, *142*, 283–303. doi: https://doi.org/10.1007/s11205-018-1910-7

Lozano, M., & Solé-Auró, A. (2021). Happiness and life expectancy by main occupational position among older workers: Who will live longer and happy? *SSM Population Health*, *13*(10735).

Luy, G., Wegner-Siegmundt, C., Wiedemann, A., & Spijker, J. (2015). Life Expectancy by Education, Income and Occupation in Germany: Estimations Using the Longitudinal Survival Method. *Comparative Population Studies*, *40*(4), 399-436. doi: https://doi.org/10.12765/CPoS-2015-16.

Luy, M., Giulio, P. D., Lego, V. D., & Sauerberg, P. L. M. (2020). Period Life Expectancy: Frequenty Used but Poorly Understood. *Gerontology*, *66*(1), 95-104. doi: 10.1159/000500955

Maccheroni, C. (2008). *Disuguaglianze nell'aspettativa di vita per grado di istruzione in Italia all'inizio degli anni 2000.* (Centre for Research on Social Dynamics "Carlo Dondena" No. 3)

Mackenbach, J., Valverde, J. R., Bopp, M., Brønnum-Hansen, H., Deboosere, P., Kalediene, R., … Nusselder, W. J. (2019). Determinants of inequalities in life expectancy: an international comparative study of eight risk factors. *The Lancet: Public Health*, *4*(10), PE529-E537. doi: https://doi.org/10.1016/S2468-2667(19)30147-1

Marmot, M. (2005). Social determinants of health inequalities. *The Lancet*, *365*, 1099–1104. doi: 10.1016/S0140-6736(05)71146-6

Marois, G., Muttarak, R., & Scherbov, S. (2020). Assessing the potential impact of COVID-19 on life expectancy. *PLOS One*, *15*(9), e0238678. doi: 10.1371/journal.pone.0238678

Marshall, V. W., Clarke, P. J., & Ballantyne, P. J. (2001, 7). Instability in the Retirement Transition. Effects on Health and Well-Being in a Canadian Study. *Research on Aging*, *23*(4), 379-409. doi: https://doi.org/10.1177/0164027501234001

Mazzaferro, C. (2019). *The Transition to NDC in Italy: Assessing Distributive and Financial Effects.* (WorkINPS Papers n. 25)

Mazzaferro, C., Morciano, M., & Savegnago, M. (2012). Differential Mortality and Redistribution in the Italian Notional Defined Contribution System. *Journal of Pension Economics and Finance*, *11*(4), 500-530.

McDonough, P., Worts, D., M.Corna, L., McMunn, A., & Sacker, A. (2017). Later-life employment trajectories and health. , *34*, 22-33.

Milligan, K., & Schirle, T. (2021). The Evolution of Longevity: Evidence from Canada. *Canadian Journal of Economics*, *Forthcoming.*. doi: https://doi.org/10.1111/caje.12497

Modig, K., Ahlbom, A., & Ebeling, M. (2021). Excess mortality from COVID-19: weekly excess death rates by age and sex for Sweden and its most affected region. *European Journal of Public Health*, *31*(1), 17-22. doi: 10.1093/eurpub/ckaa218

Moscarola, F., Fornero, E., & Strøm, S. (2016). Absenteeism, childcare and the effectiveness of pension reforms. *IZA Journal of European Labor Studies*, *5*(1). doi: https://doi.org/10.1186/s40174-016-0056-3

Muntaner, C., Solar, O., Vanroelen, C., Martinez, J., Vergara, M., Santana, V., ... EMCONET Network (2010). Unemployment, informal work, precarious employment, child labor, slavery, and health inequalities: pathways and mechanisms. *International Journal of Health Services*, *40*(2), 281-295. doi: 10.2190/HS.40.2.h

Murray, C. J. L., Lopez, A., Chin, B., Feehan, D., & Hill, K. (2006). Estimation of potential global pandemic influenza mortality on the basis of vital registry data from the 1918–20 pandemic: a quantitative analysis. *The Lancet*, *368*, 2111-2118.

Murtin, F., Mackenbach, J., Jasilionis, D., & d'Ercole, M. (2017). *Inequalities in longevity by education in OECD countries : Insights from new OECD estimates.* (OECD Statistics Working Papers, n° 2017/02)

Nielsen, J., Krause, T., & Molbak, K. (2018). Influenza-associated mortality determined from all-cause mortality, denmark 2010/11-2016/17: the flumomo model. *Influenza Other Respir Viruses*, *12*(5), 591-604. doi: 10.1111/irv.12564

Noymer, A., & Garenne, M. (2000). The 1918 Influenza Epidemic's Effects on Sex Differentials in Mortality in the United States. *Population and Development Review*, *26*(3), 565-581.

Olson, D., Simonsen, L., Edelson, P., & Morse, S. (2005). Epidemiological evidence of an early wave of the 1918 influenza pandemic in New York City. *Proceedings of the National Academy of Sciences*, *102*(31), 11059–11063. doi: 10.1073/pnas.0408290102

Pascariu, M. D., Dańko, M. J., Schöley, J., & Rizzi, S. (2018). ungroup: An R package for efficient estimation of smooth distributions from coarsely binned data. , *3, number*

= 29, journal = Journal of Open Source Software, doi =, 937.

Paterlini, M. (2020). On the front lines of coronavirus: the Italian response to covid-19. *BMJ*, *368*. doi: 10.1136/bmj.m1065.

Permanyer, I., Spijker, J., Blanes, A., & Renteria, E. (2018). Longevity and Lifespan Variation by Educational Attainment in Spain: 1960–2015. *Demography*, *55*, 2045-2070. doi: https://doi.org/10.1007/s13524-018-0718-z

Petrelli, A., Napoli, A. D., Sebastiani, G., Rossi, A., Rossi, P., Demuru, E., ... Frova, L. (2019). Atlante Italiano delle Disugualianze di Mortalità per Livello di Istruzione. *Epidemiologia e Prevenzione*, *43*(1).

Pijoan-Mas, J., & Ríos-Rull, J.-V. (2014). Heterogeneity in Expected Longevities. *Demography*, *51*(6), 2075-2102. doi: https://doi.org/10.1007/s13524-014-0346-1

Piketty, T. (2013). *Capital in the XXI Century.* Cambridge, MA: Harvard University Press.

Pirani, E., & Silvani, S. (2015). Is temporary employment damaging to health? A longitudinal study on Italian workers. *Social Science & Medicine*, *124*, 121-131. doi: https://doi.org/10.1016/j.socscimed.2014.11.033

Ponomarenko, V. (2016). Cumulative disadvantages of non-employment and non-standard work for career patterns and subjective well-being in retirement. *Advances in Life Course Research*, *30*, 133-148.

Pozzebon, S., & Mitchell, O. S. (1989). Married Women's Retirement Behavior. *Journal of Population Economics*, *1*(1), 39-53.

Principi, A., Checchucci, P., & Rosa, M. D. (2012). *Income from work after retirement in Italy.* Ancona: National Institute of Health and Science on Ageing (INRCA.

Quinlan, M. (2015). *The effects of non-standard forms of employment on worker health and safety.* International Labor Organization.

Quinlan, M., Mayhew, C., & Bohle, P. (2001). The global expansion of precarious employment, work disorganization, and consequences for occupational health: a review of recent research. *International Journal of Health Services*, *31*(2), 335-414.

Ragona, G., Giunchi, G., Rocchi, G., Muzzi, A., & Felici, A. D. (1978). Excess Mortality from Influenza in a Large Urban Population, Rome, Italy, 1956-76. *The Journal of Hygiene*, *80*(2), 249-257.

Riekhoff, A.-J. (2018). Extended working lives and late-career destabilisation: A longitudinal study of Finnish register data. *Advances in Life Course Research*, *35*, 114-125.

Riekhoff, A.-J., & Järnefelt, N. (2018). Retirement Trajectories and Income Redistribution Through the Pension System in Finland. *Social Forces*, *97*(1), 27–54.

Rizzo, C., Bella, A., Viboud, C., Simonsen, L., Rota, M., Salmaso, S., & degli Atti, M. C. (2007). Trends for Influenza-related Deaths during Pandemic and Epidemic Seasons, Italy, 1969–2001. *Emerging Infectious Diseases*, *13*(5), 694-699.

Roelfs, D. J., Shor, E., Curreli, M., Clemow, L., Burg, M. M., & Schwartz, J. E. (2012). Widowhood and Mortality: A Meta-Analysis and Meta-Regression. *Demography*, *49*(2), 575-606.

Roelfs, D. J., Shor, E., W.Davidson, K., & E.Schwartz, J. (2011, 3). Losing life and livelihood: A systematic review and meta-analysis of unemployment and all-cause mortality. *Social Science and Medicine*, *72*(6), 840-854.

Saez, E., & Zucman, G. (2016). Wealth Inequality in the United States since 1913: Evidence from Capitalized Income Tax Data. *Quarterly Journal of Economics*, *131*(2), 519-578. doi: https://doi.org/10.1093/qje/qjw004

Salm, M. (2009, 9). Does job loss cause ill health? *Health Economics*, *18*(9), 1075-1089.

Saraceno, C. (2018). Aging Women under Pressure in Italian Families. *Ethnologie française*, *3*(3), 439–450. doi: https://doi.org/10.3917/ethn.183.0439

Sarti, S., & Zella, S. (2016). Changes in the labour market and health inequalities during the years of the recent economic downturn in Italy. *Social Science Research*, *57*, 116-132.

Sasson, I. (2016). Trends in Life Expectancy and Lifespan Variation by Educational Attainment: United States, 1990–2010. *Demography*, *53*, 2692-293. doi: https://doi.org/10.1007/s13524-015-0453-7

Saunders-Hastings, P., & Krewski, D. (2016). Reviewing the History of Pandemic Influenza: Understanding Patterns of Emergence and Transmission. *Pathogens*, *5*(4), 66.

Schmitz, H. (2011). Why are the unemployed in worse health? The causal effect of unemployment on health. *Labour Economics*, *18*(1), 71-78.

Schoenbach, V. (1999). *Understanding the fundamentals of epidemiology an evolving text.* (Unpublished Manuscript)

Sen, A. (1988). Mortality as an Indicator of Economic Success and Failure. *The Economic Journal*, *108*(446), 1-25.

Serfling, R. (1953). Methods for current statistical analysis of excess pneumonia-influenza deaths. *Public Health Reports*, *78*(6), 494-506.

Siegrist, J., & Theorell, T. (2006). Soci-economic position and health. The role of work and employment. In J. Siegrist and M. Marmot (Ed.), *Social Inequalities in Health: New Evidence and Policy Implications* (p. 73-100). Oxford: Oxford University Press.

Snowden, F. N. (2019). *Epidemics and Society: From the Black Death to the Present.* New Heaven: Yale University Press.

Sullivan, D., & von Wachter, T. (2009, 8). Job Displacement and Mortality: An Analysis Using Administrative Data. *Quarterly Journal of Economics*, *124*(3), 1265–1306.

Tan, L., & Koedel, C. (2019). The Effects of Differential Income Replacement and Mortality on U.S. Social Security Redistribution. *Southern Economic Association*, *86*(2), 613-637. doi: https://doi.org/10.1002/soej.12395

Tanaka, H., Nusselder, W. J., Bopp, M., Brønnum-Hansen, H., Kaledien, R., Lee, J. S., ... Mackenbach, J. P. (2019). Mortality inequalities by occupational class among men in Japan, South Korea and eight European countries: a national register-based study. *Journal of Epidemiology and Community Health*, *73*, 750-758. doi: http://dx.doi.org/10.1136/jech-2018-211715

Thatcher, R., Kannisto, V., & Vaupel, J. (1998). *The force of mortality at ages 80 to 120.* Odense University Press.

Torssander, J., & Almquist, Y. B. (2017). Trajectories of economic, work- and health-related disadvantage and subsequent mortality risk: Findings from the 1953 Stockholm Birth Cohort. *Advances in Life Course Research*, *31*(4), 57-67.

Toulemon, L., & Barbieri, M. (2005). The Mortality Impact of the August 2003 Heat Wave in France: Investigating the 'Harvesting' Effect and Other Long-Term Consequences. *Population Studies*, *62*(1), 39-53.

Trias-Llimós, S., Riffe, T., & Bilal, U. (2020). Monitoring life expectancy levels during the COVID-19 pandemic: Example of the unequal impact of the first wave on Spanish regions. *PLOS One*, *15*(11), e0240286. doi: 10.1371/journal.pone.0241952 van der Heide, I., van Rijn, R. M., Robroek, S. J., Burdorf, A., & Proper, K. I. (2013, 9). Is retirement good for your health? A systematic review of longitudinal studies. *BMC Public Health*, *13*(9), 1180.

van Raalte, A., & Caswell, H. (2013). Perturbation Analysis of Indices of Lifespan Variability. *Demography*, *50*, 1615–1640. doi: https://doi.org/10.1007/s13524-013-0223-3

van Raalte, A., Martikainen, P., & Myrskylä, M. (2014). Lifespan Variation by Occupational Class: Compression or Stagnation Over Time? *Demography*, *51*, 73-95. doi: https://doi.org/10.1007/s13524-013-0253-x

van Raalte, A., Sasson, I., & Myrskylä, M. (2018). The case for monitoring lifespan inequality. *Science*, *362*(6418), 1002-1004. doi: https://doi.org/10.1126/science.aau5811

Vaupel, J. W. (1986). How change in age-specific mortality affects life expectancy. *Population Studies*, *40*, 147-157. doi: https://doi.org/10.1080/0032472031000141896

Vaupel, J. W., & Yashin, A. I. (1985). Heterogeneity's Ruses: Some Surprising Effects of Selection on Population Dynamics. *The American Statistician*, *39*(3), 176-185.

Vaupel, J. W., Zhang, Z., & van Raalte, A. (2011). Life expectancy and disparity: An international comparison of life table data. *BMJ Open*, *1*(e000128). doi: https://doi.org/10.1136/bmjopen-2011-000128

Viboud, C., Graisa, R. F., Lafont, B. A. P., Miller, M. A., & Simonsen, L. (2005). Multinational Impact of the 1968 Hong Kong Influenza Pandemic: Evidence for a Smoldering Pandemic. *The Journal of Infectious Diseases*, *192*(2), 233-248.

Viboud, C., Simonsen, L., Fuentes, R., Flores, J., Miller, M., & Chowell, G. (2013). Global Mortality Impact of the 1957–1959 Influenza Pandemic. *The Journal of Infectious*

Diseases, *213*, 738-745.

Visser, M., Gesthuizen, M., Kraaykamp, G., & Wolbers, M. H. J. (2016). Inequality among Older Workers in the Netherlands: A Life Course and Social Stratification Perspective on Early Retirement. *European Sociological Review*, *32*(3), 370–382. doi: https://doi.org/10.1093/esr/jcw013

Viviani, G. (2010). Il lavoro irregolare in Italia. Un'analisi longitudinale dei percorsi lavorativi. *Stato e Mercato*, *88*(1), 149-179.

Vågero, D., & Garcy, A. M. (2016). Does unemployment cause long-term mortality? Selection and causation after the 1992–96 deep Swedish recession. *The European Journal of Public Health*, *26*(5), 778-783.

Wachter, K. (2014). *Essential Demographic Methods.* Cambridge, MA: Harvard University Press.

Waenerlund, A.-K., Gustafsson, P. E., Hammarström, A., & Virtanen, P. (2014). History of labour market attachment as a determinant of health status: a 12-year follow-up of the Northern Swedish Cohort. *BMJ Open*, *4*.

Wahrendorf, M. (2015). Previous employment histories and quality of life in older ages: sequence analyses using SHARELIFE. *Ageing & Society*, *35*, 1928-1959.

Wahrendorf, M., Akinwale, B., Landy, R., Matthews, K., & Blane, D. (2017). Who in Europe Works beyond the State Pension Age and under which Conditions? Results from SHARE. *Population Ageing*, *10*, 269–285. doi: https://doi.org/10.1007/s12062-016-9160-4

Waldron, H. (2007). Trends in Mortality Differentials and Life Expectancy for Male Social Security–Covered Workers, by Socioeconomic Status. *Social Security Builletin*, *67*(3).

Weeden, K., & Grusky, D. (2012). The Three Worlds of Inequality. *American Journal of Sociology*, *117*(6), 1723-1785. doi: https://doi.org/10.1086/665035

Wenau, G., Grigoriev, P., & Shkolnikov, V. (2019). Socioeconomic disparities in life expectancy gains among retired German men, 1997–2016. *Journal of Epidemiology Community and Health*, *0*, 1-7.

Westerlund, H., Vahtera, J., Ferrie, J. E., Singh-Manoux, A., Pentti, J., Melchior, M., ... Kivima¨ki, M. (2010). Effect of Retirement on Major Chronic Conditions and Fatigue: French Gazel Occupational Cohort Study. *BMJ*, *341*, c6149.

Whitehouse, E., & Zaidi, A. (2008). *Socio-Economic Differences in Mortality: Implications for Pensions Policy.* (OECD Social, Employment and Migration Working Papers No. 71)

Wilkinson, R., & Marmot, M. (2003). *Social determinants of health. the solid facts.* World Health Organization.

Wilmoth, J., , Zureick, S., Canudas-Romo, V., Inoue, M., & Sawyer, C. (2011). A flexible two-dimensional mortality model for use in indirect estimation. *Population Studies*, *66*(1), 1-28. doi: 10.1080/00324728.2011.611411

Zang, E. (2020). Spillover effects of a husband's retirement on a woman's health: Evidence from urban China. *Social Science & Medicine*, *245*.

Zella, S., & Harper, S. (2018). The Impact of Life Course Employment and Domestic Duties on the Well-Being of Retired Women and the Social Protection Systems That Frame This. *Journal of Aging and Health*, 1-24.

Appendices

A Appendix Chapter 1

A.1 Types of pension benefits

The Italian social security provides different types of pension benefits to categories of people who are unable to work because of health reasons or reduced work capacity[16]. In the INPS LoSai sample, these benefits are grouped into three main categories:

1. **Disability pension benefits** are allowances payable to insured individuals whose work capacity is temporarily (*assegno ordinario di invalidità*) or permanently (*pensione di inabilità*) reduced due to physical or mental infirmity

2. **Indemnity pension benefits** (*pensioni indennitarie*) are payable to insured individuals upon the occurrence of an occupational injury or disease.

3. **Social pension benefits** include the following sub-categories of benefits:

    - **Social pension** *strictu senso* (*pensione/assegno sociali*): allowance payable to individuals in poor economics conditions whose income falls below a minimum threshold set by the law on a yearly basis. The right to such allowances is established on the basis of household income.

    - **Civilian disability pension** (*pensione agli invailidi civili*): allowance payable to all citizens aged 18-65 whose health conditions (including blindness, deafness, mutism) limit their work capacity completely and on a permanent basis. For individuals aged 65+, the civilian disability benefit is transformed into a permanent social allowance (*pensione/assegno sociale*, described above)

    - **Attendance allowance** (*indennità di accompagnamento*): a temporary allowance payable to people with disabilities who need permanent attendance

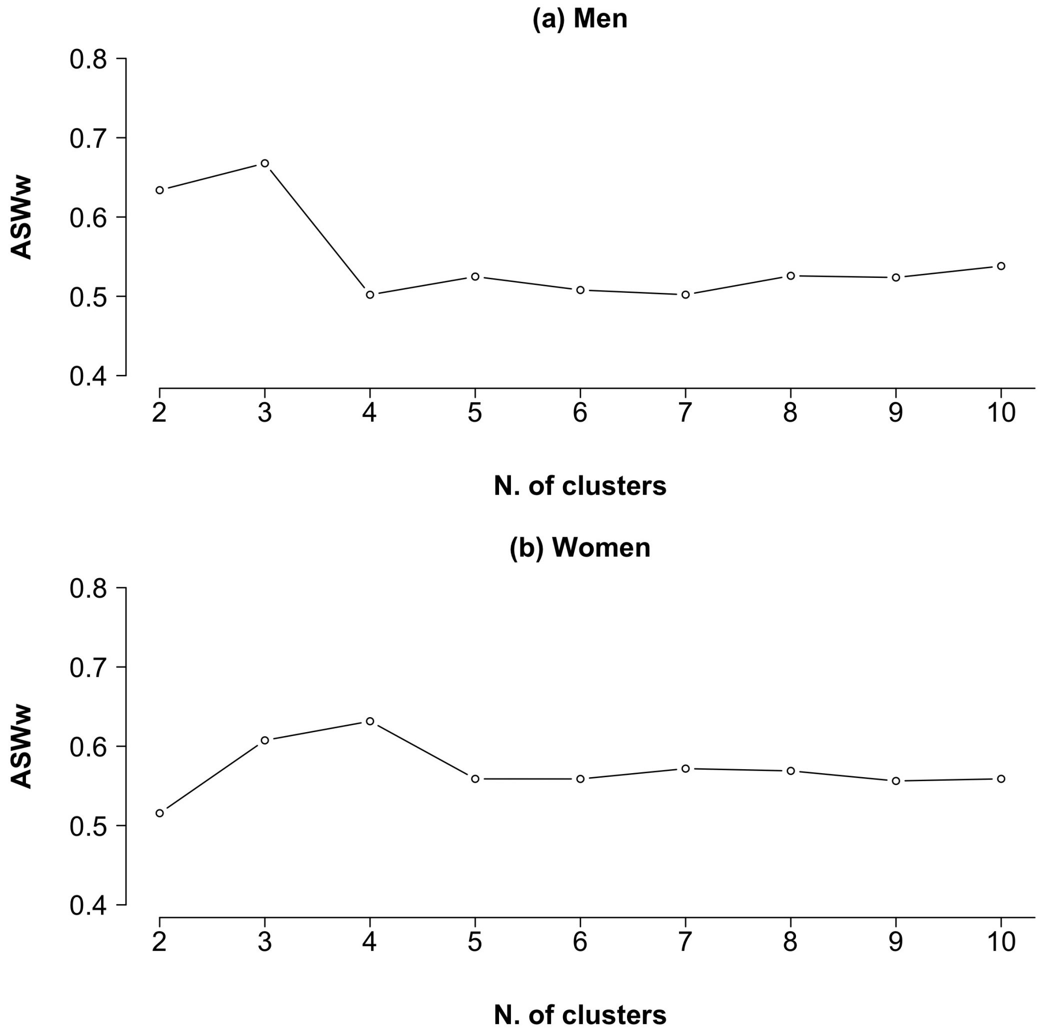
[16] Cf. `https://www.istat.it/it/files//2011/12/glossario2.pdf`

for accomplishing daily tasks and/or moving around

- **War disablement pensions** (*pensione di guerra*): allowance payable to people who have been injured or disabled as a result of any veteran service in Italian Army Forces

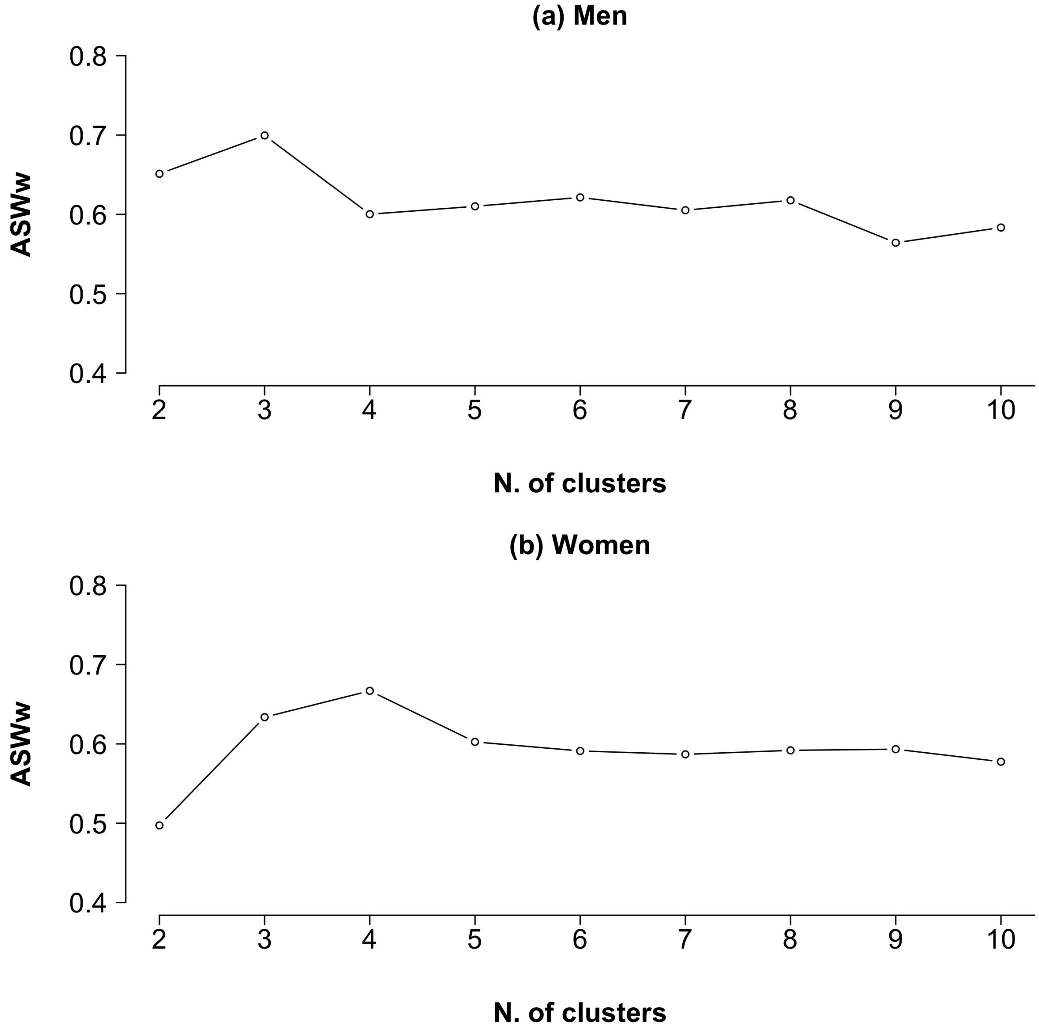
A.2 Supplementary Figures

Figure A.1: Average weighted silhouette width (ASWw) for different cluster solutions (hierarchical)



Clusters are obtained through hierarchical clustering.

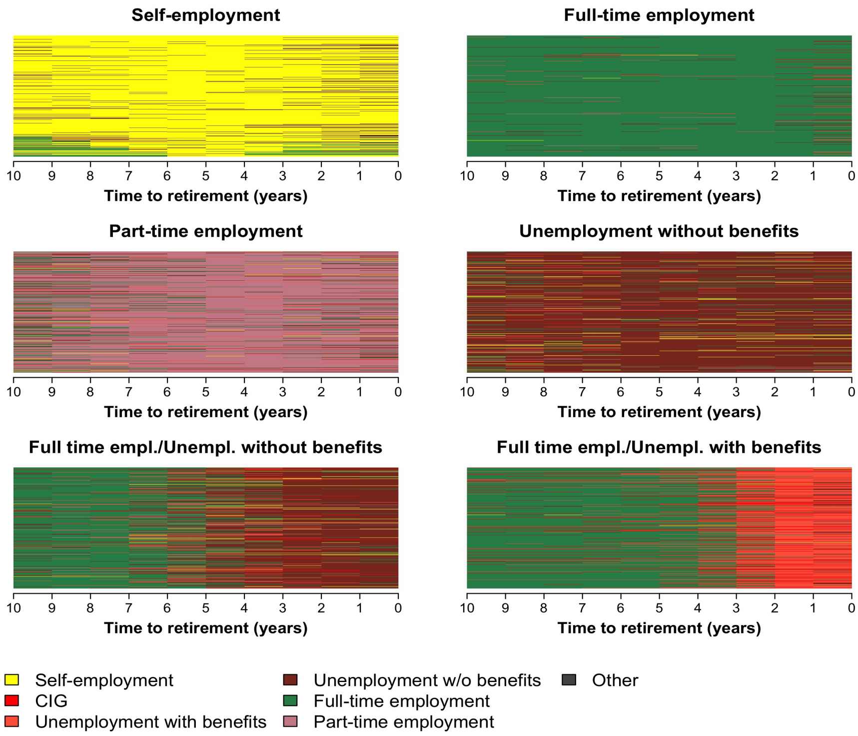
Figure A.2: Average weighted silhouette width (ASWw) for different cluster solutions (PAM)



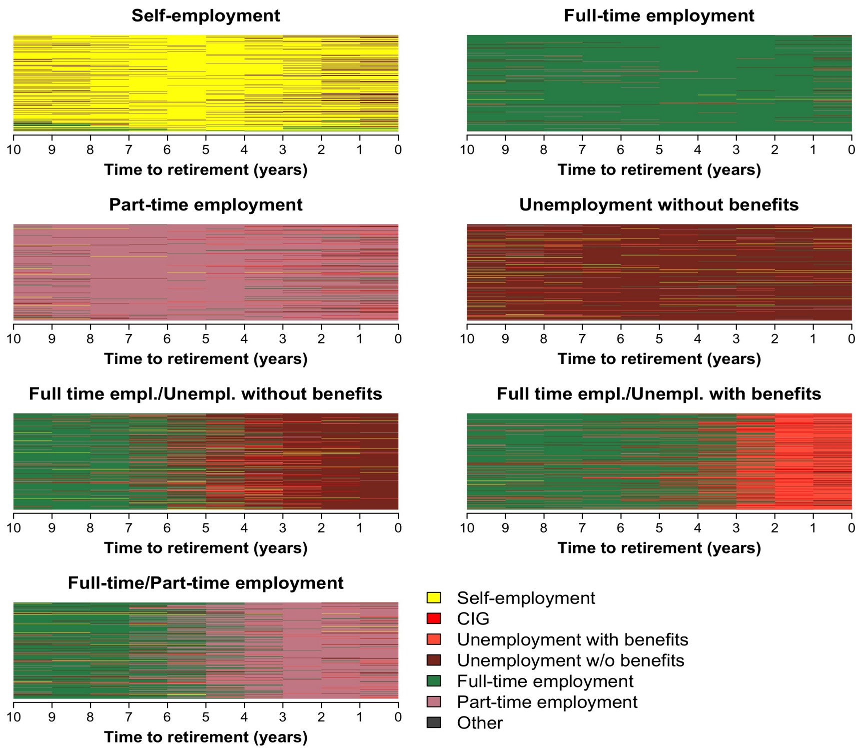
Clusters are obtained through partitioning around medoids (PAM) criterion.

Figure A.3: Clusters of ideal-type late career trajectories based on PAM clustering.

(a) Men



(b) Women



Notes. The horizontal axis measures the time to retirement (years). For each cluster, 500 representative sequences ordered based on the distance from the most frequent sequence in each cluster are shown. Own elaboration based on INPS LoSai sample.

A.3 Supplementary Tables

Table A.1: Post-retirement mortality and ideal-type late career trajectories
Men - Full results with baseline hazards

Death	(1) Model 1	(2) Model 2	(3) Model 3
Cluster (Ref: Full-time employment)			
Self-employment	0.115***	0.083**	0.090**
	(0.030)	(0.031)	(0.031)
Part-time employment	0.435***	0.165†	0.144
	(0.089)	(0.089)	(0.089)
Unemployment w/o benefits	0.404***	0.220***	0.126***
	(0.022)	(0.023)	(0.023)
Full-time empl./Unempl. w/o benefits	0.407***	0.238***	0.077*
	(0.029)	(0.029)	(0.031)
Full-time empl./Unempl. with benefits	0.105***	0.084**	0.095**
	(0.031)	(0.031)	(0.031)
Occupational status (Ref: Blue-collar)			
Manager		-0.288***	-0.237***
		(0.056)	(0.056)
White-collar		-0.211***	-0.168***
		(0.021)	(0.021)
Age at first job		-0.016***	-0.015***
		(0.003)	(0.003)
Retirement age		0.097***	0.084***
		(0.003)	(0.003)
Work after retirement		-0.525***	-0.491***
		(0.023)	(0.023)
Macroregion (Ref: Centre)			
North-East		0.070**	0.079**
		(0.026)	(0.027)
North-West		-0.003	0.011
		(0.024)	(0.024)
South		-0.123***	-0.146***
		(0.025)	(0.025)
Islands		-0.169***	-0.169***
		(0.029)	(0.029)
Disability pension			0.654***
			(0.028)
Survivor pension			0.269***
			(0.063)
Indemnity pension			0.072†
			(0.039)
Social pension			0.793***
			(0.044)
# of weeks in sickness/injury leave			-0.002
			(0.002)
# of sickess/injury leave episodes			0.010†
			(0.005)
dur1	139.401***	27.682***	40.509***
	(3.474)	(5.054)	(5.127)
dur2	142.286***	30.603***	43.429***
	(3.469)	(5.047)	(5.121)

Continued on next page

Table A.1 Continued from previous page

Death	(1) Model 1	(2) Model 2	(3) Model 3
dur3	142.417***	30.772***	43.598***
	(3.469)	(5.046)	(5.119)
dur4	142.500***	30.880***	43.707***
	(3.469)	(5.044)	(5.117)
dur5	142.518***	30.929***	43.760***
	(3.469)	(5.043)	(5.116)
dur6	142.523***	30.956***	43.792***
	(3.469)	(5.042)	(5.116)
dur7	142.682***	31.137***	43.975***
	(3.469)	(5.041)	(5.115)
dur8	142.679***	31.160***	44.001***
	(3.468)	(5.040)	(5.114)
dur9	142.767***	31.269***	44.110***
	(3.468)	(5.039)	(5.113)
dur10	142.844***	31.375***	44.216***
	(3.468)	(5.038)	(5.112)
dur11	142.940***	31.494***	44.336***
	(3.468)	(5.037)	(5.111)
dur12	142.998***	31.573***	44.413***
	(3.467)	(5.036)	(5.110)
dur13	143.061***	31.661***	44.499***
	(3.467)	(5.035)	(5.109)
dur14	143.156***	31.784***	44.621***
	(3.466)	(5.033)	(5.107)
dur15	143.147***	31.801***	44.636***
	(3.466)	(5.033)	(5.106)
dur16	143.237***	31.919***	44.754***
	(3.466)	(5.032)	(5.105)
dur17	143.251***	31.962***	44.800***
	(3.465)	(5.031)	(5.105)
dur18	143.214***	31.953***	44.806***
	(3.464)	(5.029)	(5.103)
Observations	1,431,429	1,431,429	1,431,429
Deaths	16,458	16,458	16,458

Notes. Results from complementary log-log models. Dependent variable: death occurrence (0,1).The variables *# of full weeks in sickness/injury leave* and *# of sickness/injury leave episodes* refer to ten years prior to retirement. Coefficients are not exponentiated. dur1-dur18 represent dummy variables capturing baseline hazard. Robust standard error in parentheses.
*** p<0.001, ** p<0.01, * p<0.05, † p<0.1.

Table A.2: Post-retirement mortality and ideal-type late career trajectories
Women - Full results with baseline hazards

Death	(1) Model 1	(2) Model 2	(3) Model 3
Cluster (Ref: Full-time employment)			
Self-employment	0.162*	0.190**	0.191**
	(0.068)	(0.069)	(0.069)
Part-time employment	0.096	0.016	-0.008
	(0.058)	(0.059)	(0.059)
Unemployment w/o benefits	0.253***	0.195***	0.115**
	(0.040)	(0.043)	(0.044)
Full-time/Part-time employment	0.203†	0.160	0.111
	(0.107)	(0.107)	(0.107)
Full-time empl./Unempl. w/o benefits	0.184**	0.120†	0.007
	(0.071)	(0.071)	(0.072)
Full-time empl./Unempl. with benefits	0.110	0.142*	0.170*
Year of birth	-0.035***	-0.003	-0.008
	(0.005)	(0.005)	(0.005)
Age at first job		0.010*	0.008†
		(0.005)	(0.005)
Retirewoment age		0.085***	0.079***
		(0.008)	(0.008)
Work after retirement		-0.440***	-0.404***
		(0.057)	(0.057)
Occupational status (Ref: Blue-collar)			
Manager		-0.266	-0.248
		(0.176)	(0.176)
White-collar		-0.018	0.008
		(0.035)	(0.036)
Macroregion (Ref: Centre)			
North-East		0.032	0.054
		(0.051)	(0.051)
North-West		0.103*	0.132**
		(0.047)	(0.047)
South		-0.043	-0.089
		(0.055)	(0.058)
Islands		0.021	0.003
		(0.072)	(0.072)
Disability pension			0.980***
			(0.059)
Survivor pension			0.103†
			(0.059)
Indemnity pension			-0.107
			(0.224)
Social pension			1.652***
			(0.084)
# of weeks in sickness/injury leave			-0.006*
			(0.003)
# of sickness/injury leave episodes.			0.008
			(0.008)
dur1	59.345***	-8.190	2.074
	(9.264)	(10.590)	(10.662)
dur2	61.531***	-5.989	4.277

Continued on next page

Table A.2 Continued from previous page

Death	(1) Model 1	(2) Model 2	(3) Model 3
	(9.260)	(10.583)	(10.655)
dur3	61.649***	-5.849	4.417
	(9.259)	(10.580)	(10.652)
dur4	61.838***	-5.645	4.621
	(9.258)	(10.578)	(10.650)
dur5	61.923***	-5.540	4.725
	(9.257)	(10.575)	(10.647)
dur6	61.881***	-5.569	4.696
	(9.255)	(10.573)	(10.645)
dur7	61.957***	-5.479	4.786
	(9.254)	(10.569)	(10.641)
dur8	62.085***	-5.331	4.936
	(9.255)	(10.571)	(10.642)
dur9	62.155***	-5.245	5.023
	(9.255)	(10.568)	(10.640)
dur10	62.156***	-5.225	5.045
	(9.253)	(10.566)	(10.638)
dur11	62.231***	-5.128	5.143
	(9.250)	(10.561)	(10.633)
dur12	62.361***	-4.985	5.287
	(9.249)	(10.560)	(10.631)
dur13	62.323***	-5.002	5.270
	(9.248)	(10.557)	(10.628)
dur14	62.554***	-4.750	5.521
	(9.246)	(10.554)	(10.626)
dur15	62.613***	-4.670	5.605
	(9.244)	(10.552)	(10.624)
dur16	62.730***	-4.540	5.738
	(9.243)	(10.550)	(10.621)
dur17	62.950***	-4.301	5.981
	(9.241)	(10.547)	(10.618)
dur18	62.882***	-4.354	5.939
	(9.239)	(10.546)	(10.617)
Observations	775,508	775,508	775,508
Deaths	3,921	3,921	3,921

Notes. Results from complementary log-log models. Dependent variable: death occurrence (0,1). The variables *# of full weeks in sickness/injury leave* and *# of sickness/injury leave episodes* refer to ten years prior to retirement. Coefficients are not exponentiated. dur1-dur18 represent dummy variables capturing baseline hazard. Robust standard error in parentheses.
*** p<0.001, ** p<0.01, * p<0.05, † p<0.1.

Table A.3: Percentage of individuals receiving specific types of pension benefits prior to retirement in each cluster

Cluster	Disability pension	Indemnity pension	Social pension	Survivors pension
Men				
Self-employment	3.8%	2.9%	1.2%	1.5%
Full-time employment	2.6%	4.3%	0.4%	1.2%
Part-time employment	9.1%	2.9%	3.3%	1.9%
Unemployment without benefits	12.8%	3.2%	11.0%	1.8%
Full time empl./Unempl. without benefits	21.5%	3.6%	8.9%	1.2%
Full time empl./Unempl. with benefits	2.1%	3.4%	1.6%	1.0%
Total	5.4%	3.9%	2.7%	1.3%
Women				
Cluster	Disability pension	Indemnity pension	Social pension	Survivors pension
Self-employment	2.0%	0.6%	0.4%	6.9%
Part-time employment	3.5%	0.5%	0.7%	8.1%
Full-time employment	2.0%	0.7%	0.4%	6.6%
Unemployment without benefits	6.7%	0.5%	2.7%	8.7%
Full-time/Part-time employment	5.9%	0.5%	0.5%	8.7%
Full time empl./Unempl. without benefits	9.2%	0.7%	1.6%	8.1%
Full time empl./Unempl. with benefits	1.8%	0.9%	0.6%	6.1%
Total	3.9%	0.6%	1.1%	7.4%

Table A.4: Estimated 5-, 10- and 15-year survival rate at retirement

Men			
Cluster	5-year	10-year	15-year
Full-time employment	0.9735 [0.9719-0.9752]	0.9322 [0.9283-0.9364]	0.8760 [0.8691-0.8837]
Self-employment	0.9704 [0.9680-0.9728]	0.9243 [0.9185-0.9302]	0.8620 [0.8519-0.8728]
Full-time empl./Unempl. with benefits	0.9707 [0.9683-0.9730]	0.925 [0.9193-0.9308]	0.8633 [0.8534-0.8739]
Full-time empl./Unempl. without benefits	0.9605 [0.9575-0.9637]	0.8999 [0.8928-0.9077]	0.8197 [0.8074-0.8336]
Part-time employment	0.9594 [0.9517-0.9662]	0.8972 [0.8787-0.9140]	0.8150 [0.7839-0.8444]
Unemployment without benefits	0.9606 [0.9580-0.9635]	0.9002 [0.8939-0.9071]	0.8202 [0.8094-0.8327]
Women			
Cluster	5-year	10-year	15-year
Full-time employment	0.9869 [0.9885-0.9850]	0.9654 [0.9608-0.9697]	0.9351 [0.9262-0.9433]
Self-employment	0.9846 [0.9870-0.9817]	0.9595 [0.9523-0.9658]	0.9242 [0.9106-0.9361]
Full-time empl./Unempl. with benefits	0.9854 [0.9877-0.9826]	0.9615 [0.9546-0.9675]	0.9279 [0.9150-0.9392]
Full-time empl./Unempl. without benefits	0.9842 [0.9868-0.9812]	0.9586 [0.9511-0.9652]	0.9226 [0.9084-0.9350]
Part-time employment	0.9856 [0.9877-0.9831]	0.9620 [0.9559-0.9674]	0.9288 [0.9172-0.9392]
Unemployment without benefits	0.9831 [0.9853-0.9807]	0.9557 [0.9497-0.9612]	0.9172 [0.9057-0.9279]
Full-time/Part-time empl.	0.9839 [0.9873-0.9797]	0.9578 [0.9471-0.9666]	0.9211 [0.9013-0.9375]

Notes. Cluster-specific survival rates estimated using the parameters from sex-specific baseline models (Column 1 of Table 1.4 and 1.5), setting year of birth at sex-specific mean values. 95% confidence intervals in parentheses. Survival rate in year t from retirement is calculated using the formula $S_t = exp(\sum_{k=1}^{t} log(1 - h_k))$ where h_k is the estimated complementary log-log hazard rate (see Equation **??** on p. 11).

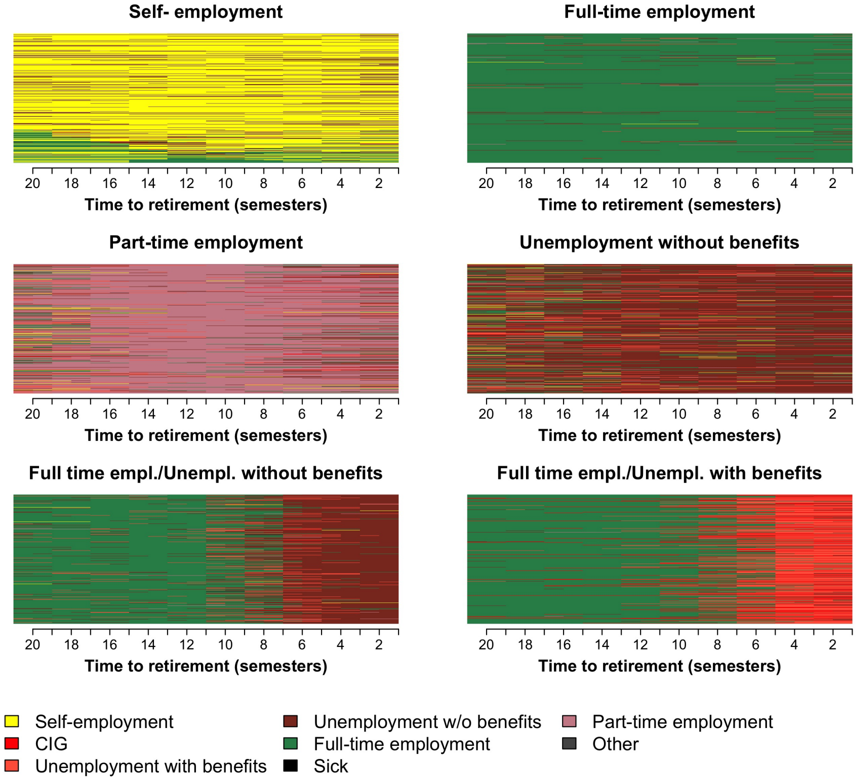
A.4 Sequence and cluster analysis based on six-month time intervals

To address concerns that coding yearly spells based on prevalent employment status over any given year masks heterogeneities in employment patterns, I construct individual sequences based on semesters (six-month periods). If an individual experiences multiple employment states during any semester, I assign him/her with the prevalent employment state over that semester, measured in terms of contributory weeks. In each semester, individuals can fall in any of the following, mutually exclusive states: (i) full-time dependent work, (ii) part-time dependent work, (iii) self-employment work, (iv) temporary suspension from work covered by wage subsidy public schemes (the so-called *Cassa Integrazione Guadagni*), (v) unemployment (i.e. having worked less than 13 weeks during a semester) covered, at least in part, by unemployment benefits, (vi) unemployment (i.e. having worked less than 13 weeks during a semester) with no unemployment benefits, (vii) sickness leave, and (viii) a residual category of statuses, as resulting from the *Estratti conto*, which do not fall into any of the former six (e.g. family leaves). It is worth noting that sickness leave is not used in year-based analysis since individuals who experience sickness leave as prevalent employment state in any of the ten years prior to retirement are excluded from the analysis. While these individuals are excluded from the semester-based analysis too, there is a minority of individuals who experience sickness leave as prevalent employment state over some semesters (not in the same year).

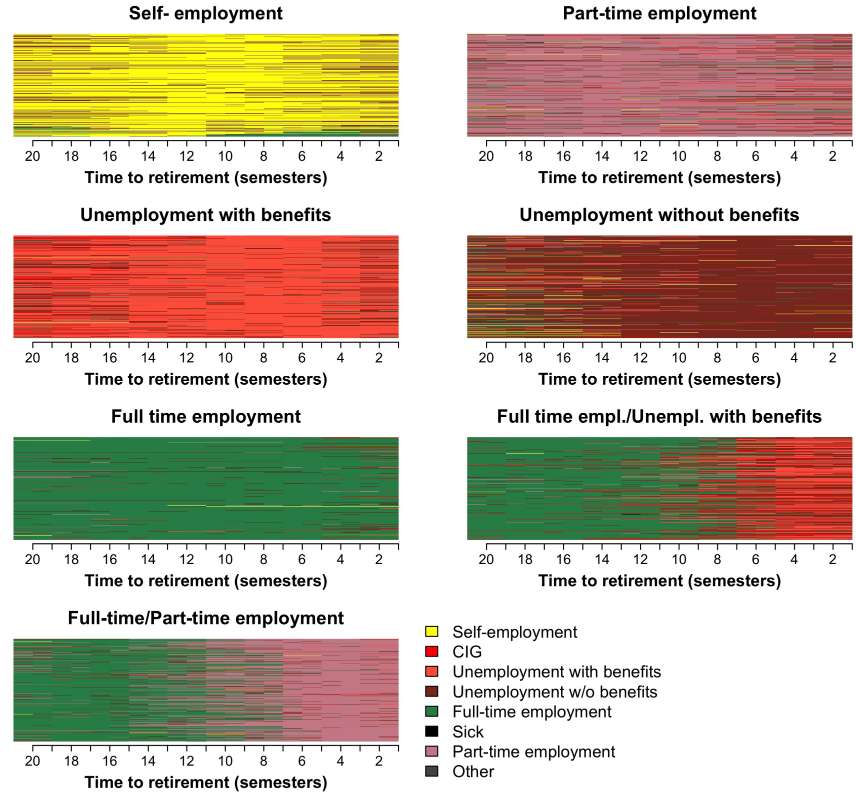
Panel (a) of Figure A.4 the composition of the semester-based six-cluster solution for men, which is virtually identical to the composition of the baseline (year-based) six-cluster solution for men. The average weighted silhouette width for the semester-based six cluster solution is slightly lower than in the year-based analysis, but still adequate (above 0.5) (Figure A.5). As shown by Table A.5, men's assignment to clusters is fairly consistent across the semester- and year-based analysis. The most noticeable discrepancy is recorded for men who, in the year-based analysis, fall in the 'Full-time employment/Unemployment without benefits' cluster. Indeed, about 39% of them fall in the same cluster when sequences are build based on semesters, about 25% fall in the 'Full-time employment cluster' and about 29% fall in the 'Unemployment without benefits' cluster. Panel (b) of Figure A.4 displays the composition of the semester-based seven-cluster solution for women. In this case, semester-based cluster analysis delivers a slightly different composition compared to the year-based analysis. Specifically, the "Full-time employment/Unemployment without benefits" cluster is replaced by a cluster dominated by sequences featuring spells of unemployment with allowances (labelled "Unemployment with benefits"). Women falling in the "Full-time employment/Unemployment without benefits" in the year-based analysis are assigned to "Full-time employment" (24%), "Full-time employment/Unemployment with benefits" (47%) and to "Unemployment without benefits" (27%) in the semester-based scenario. Assignment to other clusters is fairly consistent across the two analyses. To check whether discrepancies clusters' assignment described above matter for post-retirement mortality, I implement survival analysis using semester-based clusters as main explanatory variables, including the full set of baseline controls. Results, reported in Table A.6, confirm that going through trajectories marked by unemployment without social allowances is associated to higher post-retirement mortality risk compared to the full-time employment trajectory.

Figure A.4: Individual late-career sequences grouped by ideal-type employment trajectories

(a) Men

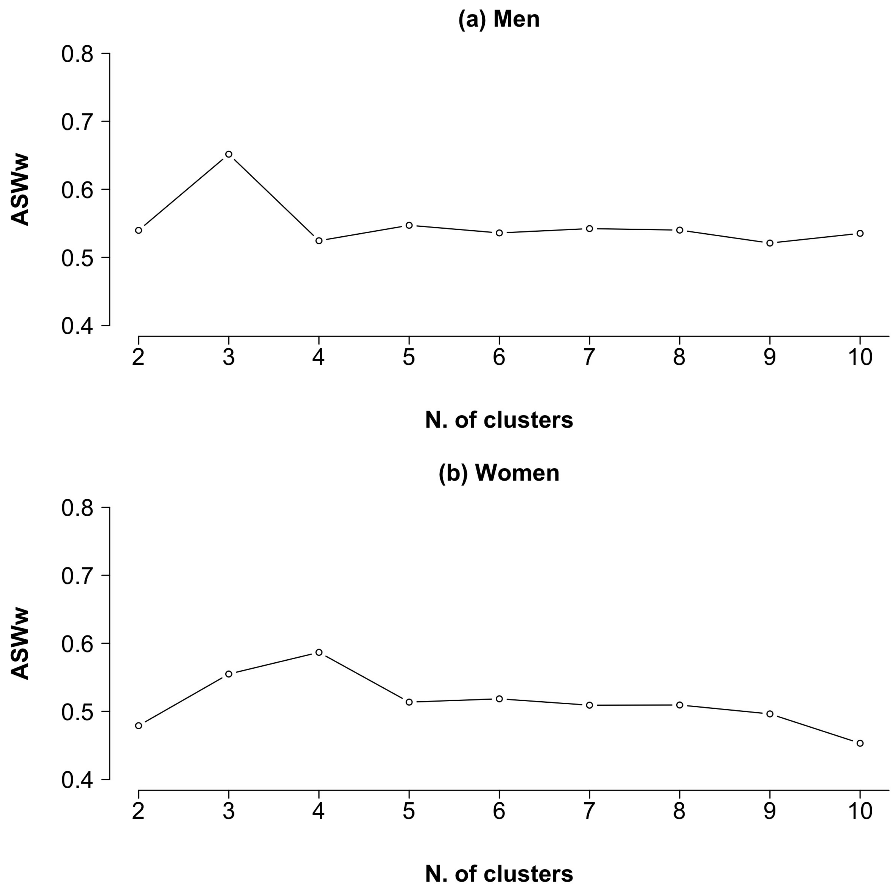


(b) Women



Clusters of ideal-type late career trajectories. For each cluster, 1000 representative sequences ordered based on the distance from the most frequent sequence in each cluster are shown. Own elaboration based on INPS LoSai sample.

Figure A.5: Average weighted silhouette width (ASWw) for different cluster solutions (semester-based sequences)



Notes. Average weighted silhouette width (ASWw) for different cluster solutions for semester-based sequences. Clusters are obtained through hierachical clustering.

Table A.5: Assignment to clusters in year-based vs semester-based approach (%)

MEN

Clusters (yearly basis)	*Clusters (six-month periods basis)* Self-empl.	Full-time empl.	Full-time empl./ Unempl. with benefits	Full-time empl./ Unempl. w/o benefits	Part-time empl.	Unempl. w/o benefits	Total
Self-empl.	99.2	0.1	0.00	0.0	0.1	0.7	100.0
Full-time empl.	2.6	95.3	1.5	0.1	0.0	0.6	100.0
Full-time empl./Unempl. with benefits	0.1	2.3	93.3	0.9	0.2	3.3	100.0
Full-time empl./Unempl. w/o benefits	2.6	25.4	4.1	38.6	0.2	29.3	100.0
Part-time empl.	3.0	7.8	0.1	0.2	82.4	6.6	100.0
Unemployment w/o benefits	12.3	14.9	0.3	1.6	1.3	69.7	100.0

WOMEN

Clusters (yearly basis)	*Clusters (six-month periods basis)* Self-empl.	Full-time empl.	Full-time/ Part-time empl.	Full-time empl./ Unempl. with benefits	Full-time empl./ Unempl. w/o benefits	Part-time empl.	Unempl. w/o benefits	Total
Self-empl.	86.7	3.9	0.0	0.2	0.7	0.2	8.4	100.0
Full-time empl.	0.2	96.1	0.3	0.7	1.6	0.1	1.00	100.0
Full-time/Part-time empl.	0.1	7.0	71.4	8.6	9.9	0.1	3.1	100.0
Full-time empl./Unempl. with benefits	0.0	7.7	0.1	80.7	0.6	8.9	2.0	100.0
Full-time empl./Unempls. w/o benefits	0.1	23.9	0.4	47.4	0.1	0.4	27.7	100.0
Part-time empl.	1.0	0.2	0.8	0.1	95.6	0.1	2.2	100.0
Unempl. w/o benefits	3.1	7.6	0.1	1.5	2.9	5.3	79.6	100.0

Notes. The table show how individuals from clusters derived in the year-based analysis distribute in clusters derived in the semester-based analysis (%). For instance, 99.2% of individuals fallin in the Self-employment cluster in the year-based analysis fall in the Self-employment cluster in the semester-based analysis.

Table A.6: Survival analysis: semester-based clusters

	(1) Men	(2) Women
Cluster (Ref: Full-time empl.)		
Self-employment	1.063*	1.188*
	(0.028)	(0.080)
Part-time employment	1.075	0.988
	(0.091)	(0.055)
Unempl. w/o benefits	1.105***	1.096*
	(0.026)	(0.047)
Full-time empl./Unempl. with benefits	1.080*	1.052
	(0.033)	(0.065)
Full-time empl./Unempl. w/o benefits	1.046	
	(0.045)	
Full-time/Part-time empl.		0.968
		(0.121)
Unempl. with benefits		1.040
		(0.111)
Year of birth	0.973***	0.993
	(0.002)	(0.005)
Manager	0.787***	0.779
	(0.044)	(0.137)
White-collar	0.839***	0.999
	(0.018)	(0.035)
Age first job	0.985***	1.008†
	(0.003)	(0.005)
Retirement age	1.089***	1.081***
	(0.004)	(0.009)
Work after retirement	0.612***	0.667***
	(0.014)	(0.038)
Macro-region of residence (Ref: Centre)		
North-East	1.088**	1.057
	(0.029)	(0.054)
North-West	1.014	1.144**
	(0.024)	(0.054)
South	0.869***	0.900†
	(0.021)	(0.051)
Islands	0.853***	0.995
	(0.025)	(0.072)
Invalidity pension	1.934***	2.609***
	(0.053)	(0.152)
Survivor pension	1.312***	1.109†
	(0.083)	(0.066)
Indemnity pension	1.073†	0.895
	(0.042)	(0.201)
Social pension	2.215***	5.160***
	(0.099)	(0.431)
Observations	1,431,429	775,508
Deaths	16,458	3,921

Notes. Results from complementary log-log models. Dependent variable: death occurrence (0,1). All models include 18 duration dummies (baseline hazard). Coefficients are expressed in the exponentiated form (hazard ratios). Robust standard error in parentheses.
*** $p<0.001$, ** $p<0.01$, * $p<0.05$, † $p<0.1$.

B Appendix Chapter 2

Indirect standardization

Since the number of deaths observed in each age group for most occupational groups is small, we adjust occupation- and sex-specific mortality rates by age through the indirect method. Indirect standardization is more appropriate than direct standardization when stratum-specific rates in the study population are unstable and susceptible to be strongly influenced by random variability (Schoenbach 1999). In indirect standardization one takes stratum-specific rates from a standard population of sufficient size. Such rates are then averaged using stratum-specific weights provided by the study population and used to calculate the expected number of deaths in the study population. Indirect adjusted mortality rate in the study population is obtained by multiplying crude-death rate in the standard population by the Standardized Mortality Ratio (SMR), given by the ratio between observed and expected deaths across all strata in the study population.

In our case, occupational groups for men and women represent the study populations. We consider two strata, i.e. the age bands 65-69 and 70-74, and we use sex-age-specific rates across all occupations as standard. For each occupation, the total number of expected deaths for each sex s is computed as:

$$\text{Expected deaths}^{o,s} = \sum R_j^s \times w_j^{o,s} \tag{2}$$

where R_j^s are sex-specific standard death rates in age band j and $w_j^{o,s}$ are age-specific weights in the chosen occupation o for each sex s. Standardized Mortality Ratios (SMR) are computed as:

$$SMR^{o,s} = \frac{\text{Expected deaths}^{o,s}}{\text{Observed deaths}^{o,s}} \tag{3}$$

where the numerator represents the total number of observed deaths in occupation o, for each sex. Finally, sex-specific indirect adjusted mortality rates for occupation o are computed as:

$$IMR^{o,s} = SMR^{o,s} \times C^s \tag{4}$$

where C^s is the sex-specific standard crude death rate.

Supplementary Tables

Table B.1: Test of Proportionality of Hazards Assumption
Men (macro-occupational groups)

Variable	ρ	χ^2	df	Prob> χ^2
Managers	-0.00185	0.05	1	0.8237
Professionals	0.00348	0.17	1	0.6795
Technicians	-0.00600	0.52	1	0.4708
Service and sales workers	0.00369	0.20	1	0.6583
Craft and related trade workers, skilled agricultural, forestry and fishery workers	-0.00393	0.22	1	0.6357
Plant and machine operators, assemblers	-0.00653	0.62	1	0.4319
Elementary occupations	-0.01208	2.12	1	0.1456
Widow	0.00664	0.64	1	0.4240
Separated/Divorced	0.00120	0.02	1	0.8853
Never married	0.01152	1.96	1	0.1611
Abroad	-0.01087	1.69	1	0.1938
Islands	0.00218	0.07	1	0.7941
North-East	-0.01661	3.97	1	0.0464
North-West	-0.00944	1.29	1	0.2564
South	0.00519	0.39	1	0.5346
Social disability pension	-0.06081	69.19	1	0.0000
Disability pension	-0.01984	6.06	1	0.0138
Secondary education	0.01522	3.38	1	0.0662
Tertiary education	0.00679	0.65	1	0.4206
Global test		99.17	19	0.0000

*** p<0.001, ** p<0.01, * p<0.05, † p<0.1.

Table B.2: Test of Proportionality of Hazards Assumption
Women (macro-occupational groups)

Variable	ρ	χ^2	df	Prob> χ^2
Managers	0.00207	0.02	1	0.8843
Professionals	0.01618	1.29	1	0.2562
Technicians	0.00986	0.49	1	0.4843
Service and sales workers	0.02281	2.67	1	0.1023
Craft and related trade workers, skilled agricultural, forestry and fishery workers	0.00683	0.24	1	0.6260
Plant and machine operators, assemblers	0.02370	2.80	1	0.0945
Elementary occupations	-0.00115	0.01	1	0.9353
Widow	-0.01494	1.14	1	0.2858
Separated/Divorced	0.02016	2.14	1	0.1437
Never married	0.02055	2.28	1	0.1314
Abroad	-0.02862	4.01	1	0.0453
Islands	0.03470	5.95	1	0.0147
North-East	-0.00239	0.03	1	0.8672
North-West	-0.00685	0.24	1	0.6267
South	0.02274	2.62	1	0.1058
Social disability pension	-0.12164	110.78	1	0.0000
Disability pension	-0.01641	1.45	1	0.2288
Secondary education	0.01154	0.66	1	0.4163
Tertiary education	0.00650	0.20	1	0.6508
Global test		139.04	19	0.0000

*** $p<0.001$, ** $p<0.01$, * $p<0.05$, † $p<0.1$.

Table B.3: Test of Proportionality of Hazards Assumption
Men (micro-occupational groups)

Variable	ρ	χ^2	df	Prob> χ^2
Legislators and senior officials	-0.00332	0.16	1	0.6883
Managing directors and chief executives	0.00008	0.00	1	0.9924
Professional services managers	-0.01163	1.99	1	0.1584
Science professionals (mathematics, computer science, chemistry, physics, biology)	-0.00295	0.12	1	0.7243
Engineers, architects and similar professions	-0.00309	0.14	1	0.7113
Life science professionals	0.01221	2.11	1	0.1460
Health professionals	-0.00707	0.71	1	0.4004
Legal, social and cultural professionals	0.00877	1.10	1	0.2947
Teaching and research professionals	-0.00580	0.48	1	0.4898
Science and engineering technicians	-0.01044	1.57	1	0.2105
Life science technicians	-0.01065	1.64	1	0.2009
Business and administration technicians	0.00592	0.50	1	0.4775
Public service technicians	-0.01247	2.25	1	0.1339
Customer service clerks	0.00699	0.71	1	0.4004
Other clerical support workers	-0.00709	0.73	1	0.3916
Numerical and material recording clerks	-0.00956	1.34	1	0.2476
Sales workers	-0.01693	4.12	1	0.0423
Personal service workers	0.01120	1.82	1	0.1773
Personal care workers	0.01159	1.93	1	0.1645
Protective service workers	0.00929	1.24	1	0.2652
Mining, building and related trade workers	-0.00189	0.05	1	0.8199
Electrical and electronic trades workers	0.00249	0.09	1	0.7640
Handicraft and printing workers	0.00261	0.10	1	0.7546
Skilled agricultural, forestry and fishery workers	-0.00842	1.02	1	0.3125
Food processing, wood working, garment & other craft & related trades workers	-0.00960	1.34	1	0.2470
Stationary plant operators	-0.01121	1.81	1	0.1788
Assemblers	-0.01008	1.47	1	0.2256
Machine operators in agricultural/food industry, drivers & mobile plant operators	-0.00087	0.01	1	0.9168
Unskilled sales workers, cleaners and helpers	-0.00788	0.90	1	0.3425
Unskilled agricultural, forestry and fishery workers	-0.00274	0.11	1	0.7425
Labourers in mining, construction, manufacturing	-0.01314	2.49	1	0.1143
Widow	0.00626	0.57	1	0.4509
Separated/Divorced	0.00084	0.01	1	0.9189
Never married	0.01164	2.01	1	0.1562
Abroad	-0.01122	1.80	1	0.1797
Islands	0.00250	0.09	1	0.7647
North-East	-0.01595	3.66	1	0.0556
North-West	-0.00867	1.09	1	0.2969
South	0.00664	0.63	1	0.4270
Social disability pension	-0.06133	70.56	1	0.0000
Disability pension	-0.01991	6.15	1	0.0132
Secondary education	0.01513	3.34	1	0.0678
Tertiary education	0.00968	1.31	1	0.2517
Global Test		132.17	43	0.0000

*** p<0.001, ** p<0.01, * p<0.05, † p<0.1.

Table B.4: Test of Proportionality of Hazards Assumption
Women (micro-occupational groups)

Variable	ρ	χ^2	df	Prob> χ^2
Legislators and senior officials	0.00463	0.10	1	0.7473
Managers	0.00187	0.02	1	0.8957
Science, engineers, architects, life science, health professionals	0.01954	1.94	1	0.1632
Legal, social and cultural professionals	0.01657	1.36	1	0.2440
Teaching and research professionals	0.01175	0.67	1	0.4131
Science, engineering, life science technicians	0.00279	0.04	1	0.8418
Business and administration technicians	0.01342	0.90	1	0.3427
Public service technicians	0.01442	1.03	1	0.3109
Customer service clerks	-0.00648	0.22	1	0.6367
Other clerical support workers	0.02212	2.43	1	0.1193
Sales workers	0.02230	2.53	1	0.1119
Personal service workers	0.01767	1.56	1	0.2119
Personal care workers	0.02191	2.43	1	0.1187
Protective service workers	-0.00202	0.02	1	0.8864
Mining, building and related trade workers	0.01745	1.60	1	0.2066
Other craft and related trade workers	-0.00354	0.06	1	0.8040
Plant and machine operators	0.01385	0.92	1	0.3364
Assemblers	0.02131	2.28	1	0.1312
Unskilled sales workers	-0.00652	0.21	1	0.6445
Cleaners and helpers	-0.00889	0.40	1	0.5253
Unskilled workers in agriculture, forestry, fishery, mining, construction, manufacturing	0.02279	2.54	1	0.1112
Widow	-0.01479	1.12	1	0.2900
Separated/Divorced	0.01987	2.08	1	0.1492
Never married	0.02174	2.57	1	0.1089
Abroad	-0.02412	2.90	1	0.0887
Islands	0.03409	5.74	1	0.0166
North-East	-0.00284	0.04	1	0.8425
North-West	-0.00737	0.27	1	0.6005
South	0.02288	2.63	1	0.1050
Social disability pension	-0.12261	112.83	1	0.0000
Disability pension	-0.01605	1.38	1	0.2406
Secondary education	0.00862	0.37	1	0.5434
Tertiary education	0.00377	0.07	1	0.7942
Global test		151.22	33	0.0000

*** p<0.001, ** p<0.01, * p<0.05, † p<0.1.

Table B.5: Partial life expectancy 65-74
Men (macro-occupational groups)

Occupational class	e_{65-74}	CI_{low}	CI_{up}
Managers and senior officials	8.69	8.72	8.65
Professionals	8.65	8.67	8.63
Technicians	8.56	8.59	8.54
Clerical support workers	8.47	8.50	8.44
Service and sales workers	8.44	8.49	8.39
Craft & related trade workers/skilled, agricultural, forestry & fishery workers	8.41	8.44	8.37
Plant and machine operators, assemblers	8.38	8.42	8.34
Elementary occupations	8.38	8.41	8.35

Table B.6: Life expectancy at 65
Men (macro-occupational groups)

Occupational class	e_{65-74}	CI_{low}	CI_{up}
Managers and senior officials	20.27	20.34	20.20
Professionals	19.14	19.19	19.10
Technicians	18.53	18.67	18.38
Clerical support workers	16.84	17.01	16.68
Service and sales workers	16.90	17.09	16.70
Craft & related trade workers/skilled agricultural, forestry & fishery workers	16.89	17.02	16.75
Plant and machine operators, assemblers	16.82	17.10	16.55
Elementary occupations	16.87	16.95	16.79

Table B.7: Partial life expectancy 65-74
Men (micro-occupational groups)

Occupational class	e_{65-74}	CI_{low}	CI_{up}
Legislators and senior officials	8.70	8.65	8.74
Managing directors and chief executives	8.68	8.63	8.73
Professional services managers	8.64	8.47	8.75
Science professionals (mathematics, computer science, chemistry, physics, biology)	8.65	8.52	8.74
Engineers. architects and similar professions	8.73	8.62	8.81
Life science professionals	8.66	8.52	8.75
Health professionals	8.68	8.63	8.72
Legal, social and cultural professionals	8.60	8.57	8.64
Teaching and research professionals	8.68	8.65	8.71
Science and engineering technicians	8.56	8.51	8.60
Life science technicians	8.55	8.49	8.60
Business and administration technicians	8.57	8.53	8.60
Public service technicians	8.56	8.48	8.64
General and keyboard clerks	8.48	8.44	8.51
Customer service clerks	8.47	8.38	8.55
Other clerical support workers	8.47	8.33	8.58
Numerical and material recording clerks	8.32	8.15	8.46
Sales workers	8.45	8.37	8.52
Personal service workers	8.36	8.24	8.47
Personal care workers	8.46	8.25	8.61
Protective service workers	8.46	8.37	8.54
Mining. building and related trade workers	8.45	8.40	8.49
Electrical and electronic trades workers	8.36	8.30	8.42
Handicraft and printing workers	8.43	8.25	8.57
Skilled agricultural, forestry and fishery workers	8.40	8.28	8.50
Food processing, wood working, garment and other craft and related trades workers	8.37	8.27	8.46
Stationary plant operators	8.45	8.35	8.53
Assemblers	8.33	8.24	8.41
Machine operators in agricultural/forestry, drivers & mobile plant operators	8.38	8.33	8.43
Unskilled sales workers, cleaners and helpers	8.37	8.32	8.41
Unskilled agricultural, forestry and fishery workers	8.48	8.40	8.56
Labourers in mining, construction, manufacturing	8.35	8.26	8.42

Table B.8: Life expectancy at 65
Men (micro-occupational groups)

Occupational class	e_{65-74}	CI$_{low}$	CI$_{up}$
Legislators and senior officials	20.36	20.30	20.44
Managing directors and chief executives	20.23	20.46	19.97
Professional services managers	19.64	20.48	18.63
Science professionals (mathematics, computer science, chemistry, physics, biology)	19.79	20.90	18.57
Engineers, architects and similar professions	20.75	21.10	20.30
Life science professionals	16.78	16.29	17.40
Health professionals	19.83	19.87	19.78
Legal, social and cultural professionals	18.40	18.51	18.27
Teaching and research professionals	19.50	19.50	19.51
Science and engineering technicians	18.82	19.19	18.43
Life science technicians	18.69	19.02	18.33
Business and administration technicians	18.16	18.34	17.98
Public service technicians	19.11	19.20	19.00
General and keyboard clerks	16.92	17.09	16.75
Customer service clerks	16.40	16.83	15.93
Other clerifcal support workers	16.59	17.61	15.51
Numerical and material recording clerks	16.26	17.45	15.02
Sales workers	18.16	18.58	17.70
Personal service workers	16.25	16.50	15.95
Personal care workers	17.62	18.12	16.97
Protective service workers	15.73	15.97	15.46
Mining, building and related trade workers	17.16	17.22	17.08
Electrical and electronic trades workers	16.18	16.57	15.78
Handicraft and printing workers	17.38	18.55	16.11
Skilled agricultural, forestry and fishery workers	17.82	18.10	17.50
Food processing, wood working, garment and other craft and related trades workers	16.67	17.00	16.30
Stationary plant operators	18.88	19.77	17.96
Assemblers	16.56	17.20	15.90
Machine operators in agricultural/food industry, drivers & mobile plant operators	16.38	16.66	16.09
Unskilled sales workers, cleaners and helpers	16.76	16.88	16.63
Unskilled agricultural, forestry and fishery workers	17.60	17.51	17.71
Labourers in mining, construction, manufacturing	16.45	16.72	16.15

Table B.9: Partial life expectancy 65-74
Women (macro-occupational groups)

Occupational class	e_{65-74}	CI$_{low}$	CI$_{up}$
Managers and senior officials	8.74	8.80	8.67
Professionals	8.78	8.79	8.76
Technicians	8.75	8.77	8.73
Clerical support workers	8.71	8.74	8.68
Service and sales workers	8.72	8.76	8.69
Craft & related trade workers/skilled agricultural, forestry & fishery workers	8.73	8.77	8.67
Plant and machine operators, assemblers	8.66	8.75	8.53
Elementary occupations	8.72	8.74	8.69

Table B.10: Life expectancy at 65
Women (macro-occupational groups)

Occupational class	e_{65-74}	CI$_{low}$	CI$_{up}$
Managers	21.83	22.41	21.19
Professionals	21.67	21.87	21.47
Technicians	21.37	21.53	21.20
Clerical support workers	20.95	21.37	20.53
Service and sales workers	20.29	20.63	19.94
Craft & related trade workers/skilled agricultural, forestry & fishery workers	21.48	22.30	20.62
Plant and machine operators, assemblers	19.39	21.47	17.37
Elementary occupations	21.39	21.53	21.23

Table B.11: Partial life expectancy 65-74
Women (micro-occupational groups)

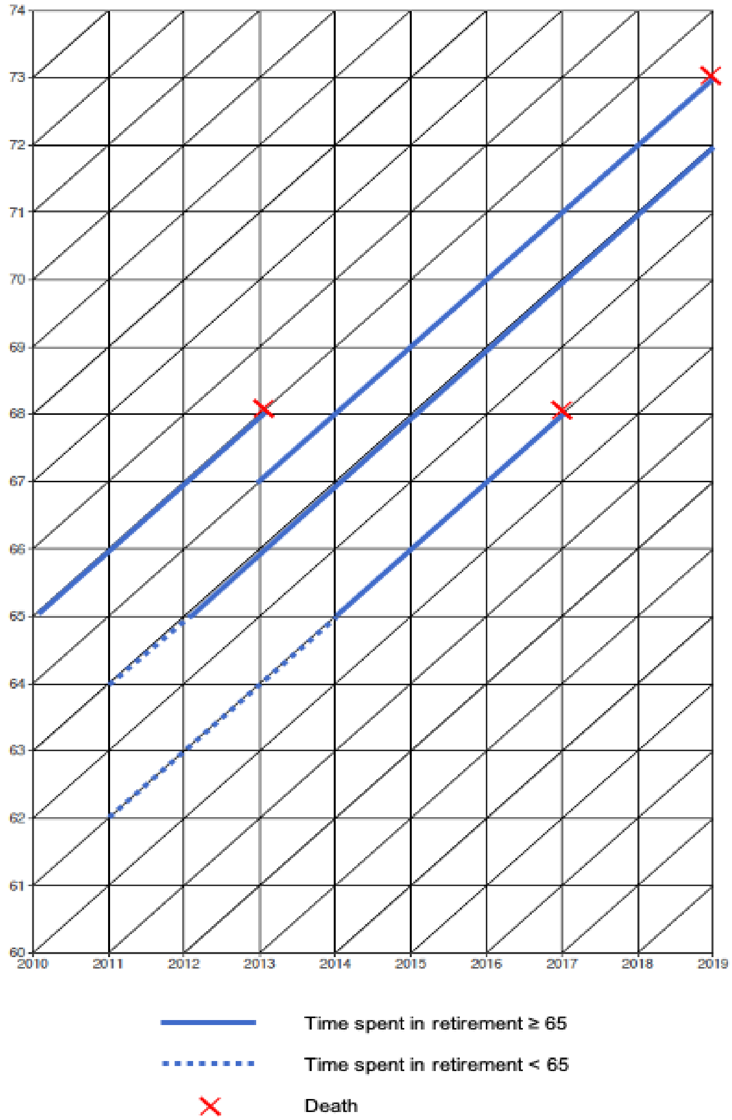
Occupational class	e_{65-74}	CI_{low}	CI_{up}
Legislators and senior officials	8.77	8.83	8.68
Managers	8.68	8.78	8.53
Legal social and cultural professionals	8.73	8.77	0.87
Teaching and research professionals	8.79	8.81	8.77
Other professionals	8.78	8.85	8.69
Science. engineering. life science technicians	8.72	8.76	8.67
Business and administration technicians	8.73	8.77	8.69
Public service technicians	8.78	8.81	8.75
General and keyboard clerks	8.72	8.75	8.69
Customer service clerks	8.70	8.78	8.59
Other clerical support workers	8.63	8.75	8.46
Sales workers	8.68	8.76	8.59
Personal service workers	8.73	8.79	8.65
Personal care workers	8.72	8.81	8.57
Protective service workers	8.75	8.80	8.68
Mining, building and related trade workers	8.71	8.78	8.64
Other craft and related trade workers	8.75	8.81	8.67
Plant and machine operators	8.68	8.83	8.37
Assemblers	8.66	8.76	8.50
Unskilled sales workers	8.69	8.73	8.65
Cleaners and helpers	8.76	8.81	8.69
Unskilled workers in agriculture. forestry. fishery. mining. construction. manufacturing	8.75	8.80	8.69

Table B.12: Life expectancy at 65
Women (micro-occupational groups)

Occupational class	e_{65-74}	CI_{low}	CI_{up}
Legislators and senior officials	22.17	22.72	21.53
Managers	20.81	22.06	19.40
Legal, social and cultural professionals	20.46	20.83	20.06
Teaching and research professionals	22.11	22.34	21.88
Other professionals	21.28	22.57	19.89
Science, engineering, life science technicians	20.90	21.52	20.26
Business and administration technicians	21.32	21.70	20.92
Public service technicians	21.03	20.91	21.17
General and keyboard clerks	21.32	21.75	20.87
Customer service clerks	19.30	20.73	17.81
Other clerical support workers	18.49	20.70	16.26
Sales workers	19.04	19.96	18.06
Personal service workers	20.56	21.18	19.88
Personal care workers	18.57	20.02	16.97
Protective service workers	21.66	21.89	21.39
Mining, building and related trade workers	21.19	21.93	20.38
Other craft and related trade workers	22.80	24.58	20.99
Plant and machine operators	19.99	23.66	16.26
Assemblers	19.29	21.68	16.98
Unskilled sales workers	21.21	21.51	20.88
Cleaners and helpers	21.99	22.36	21.56
Unskilled workers in agriculture, forestry, fishery, mining, construction, manufacturing	21.71	21.69	21.73

Supplementary Figures

Figure B.1: Lexis-type diagram



C Appendix Chapter 3

C.1 Reforming the Italian pension system: an overview

Since the mid-1970s, the Italian population has been ageing fast. Between 1950 and 2018, the percentage of individuals aged 65 or more has increased from 9.5% to 22.7% of the total population (World Bank, 2019). Initially, policymakers neglected the implications of the ongoing demographic shift for the financial sustainability of the pay-as-you-go pension system. Indeed, the growth of working age population, which followed the baby boom of the early 1960s, and high employment levels ensured enough revenues for covering welfare outlays. As the first signals of rapidly rising pension expenditure began to appear at the beginning of the 1990s, policymakers inaugurated a long series of pension reforms aimed at extending the length of working life and reducing pension disbursements. In 1992, the Amato reform (Law n. 503/1992), while maintaining defined benefit pension rules, introduced three major changes (i) it increased progressively legal retirement age, up to 60 for women and 65 for men; (ii) it increased the number of years over which pensionable earnings were to be computed; (iii) it modified the indexation mechanism linking the growth of pension benefits to price inflation in lieu of real earnings growth. In 1995, the Dini reform (Law n. 335/1995) determined the transition from DB to NDC pension rules, with the aim of tightening the link between pension benefits and contributions. The phase-in period was set to be very gradual. Workers with at least 18 years of contributions as of December 1995 were fully unaffected by the reform. Instead, those with a shorter contributory record were to be affected on a *pro rata* basis, the weight of DB depending on the ratio between pre-1995 to the overall contribution period upon retirement. In addition, the 1995 reform tightened age requirement for accessing seniority pension benefits. Further tightening of age requirements for claiming seniority pension benefits was also at the core of the Maroni Reform in 2004 and the Prodi reform in 2007. In 2011, the Fornero reform (Law Decree n. 201/2011) accelerated the transition to full NDC rules, introducing a *pro rata* contribution for all workers starting from January

1, 2012. This means that all pensions awarded from this date onward have an NDC component, regardless of the 18-year contribution period mentioned above. The Fornero Reform provided also for (i) the abolition of seniority pension, which was replaced by the so-called "anticipated" pension, (ii) the gradual convergence towards a unique longevity-indexed retirement age, independent of gender and occupational profile, set to reach 67 on January 1, 2019, and (ii) the automatic update of minimum retirement age, and related conversion factors, every two years from 2019 onward. After the Fornero reform, major changes gave way to experimental and temporary measures which aimed at providing more flexibility in the retirement timing.[17]

Supplementary Tables

Table C.1: Obs. per individual between ages 45-49 in the final *Dichiarazioni Uniemens* sample

Obs per individual	Women	Men	Total
1	261,057	398,564	659,621
	12.4%	8.2%	9.5%
2	187,852	295,593	782,808
	8.9%	6.1%	7.0%
3	161,069	272,062	433,131
	7.5%	5.9%	6.4%
4	174,332	341,581	515,913
	8.3%	7.1%	7.4%
5	1,322,630	3,534,506	4,857,136
	62.8%	73.0%	69.9%
Total	2,106,940	4,842,306	6,949,246
	100%	100%	100%

Notes. Own elaboration on INPS data.

Table C.2: *Dichiarazioni Uniemens* sample
Observations by sex and prevalent occupation position

Sex	Blue-collar	White-collar	Managers	Total
Women	1,281,863	709,035	15,316	2,006,214
Men	3,321,301	1,243,378	157,306	4,721,985
Total	4,603,164	1,952,413	172,622	6,728,199

[17]For instance, Law n. 232/2016 introduced the possibility for individuals aged 63 or more to claim a specific social allowance until the attainment of legal retirement age/fulfillment of requirements for claiming anticipated pension benefits. Law Decree No. 4/2019 introduced an anticipated retirement option, valid for the 2019-2021 triennium, for workers aged 62 with 38 years of contributions (so-called *quota 100*).

Table C.3: *Dichiarazioni Uniemens*
Observations by year of birth and prevalent occupation
Men

Year of birth	Blue-collar	White-collar	Managers	Total
1930	147,407	32,777	3,116	183,3
1931	141,972	32,998	3,343	178,313
1932	136,486	33,438	3,564	173,488
1933	138,791	35,064	3,845	177,7
1934	134,759	35,868	4,176	174,803
1935	131,979	38,31	4,442	174,731
1936	121,986	37,823	4,516	164,325
1937	123,278	40,028	5,289	168,595
1938	124,128	44,002	5,911	174,041
1939	118,874	43,633	5,806	168,313
1940	116,921	44,367	6,478	167,766
1941	101,482	41,563	6,634	149,679
1942	94,447	40,247	6,405	141,099
1943	91,963	41,226	6,619	139,808
1944	94,543	42,444	5,9	142,887
1945	88,952	41,042	5,669	135,663
1946	112,945	56,155	7,644	176,744
1947	113,023	56,167	7,313	176,503
1948	115,732	57,02	6,979	179,731
1949	112,985	52,923	7,42	173,328
1950	113,083	50,354	7,001	170,438
1951	109,944	50,71	6,772	167,426
1952	110,746	49,493	6,514	166,753
1953	113,959	48,309	6,283	168,551
1954	121,697	48,104	4,905	174,706
1955	125,401	48,531	4,876	178,808
1956	129,292	50,167	4,881	184,34
1957	134,539	50,623	4,986	190,148
Total	3,321,301	1,243,378	157,306	4,721,985

Table C.4: *Dichiarazioni Uniemens*
Observations by year of birth and prevalent occupation
Women

Year of birth	Blue-collar	White-collar	Managers	Total
1930	45,675	13,192	131	58,998
1931	44,858	13,326	139	58,323
1932	44,764	13,465	161	58,39
1933	45,163	14,078	172	59,413
1934	46,193	14,524	181	60,898
1935	46,717	15,554	208	62,479
1936	44,042	15,769	203	60,014
1937	44,952	16,817	263	62,032
1938	44,975	18,301	345	63,621
1939	43,027	18,562	302	61,891
1940	43,82	20,014	355	64,189
1941	39,141	19,69	333	59,164
1942	35,448	18,473	366	54,287
1943	36,364	19,429	477	56,27
1944	38,104	20,864	417	59,385
1945	35,56	20,393	435	56,388
1946	46,184	28,408	634	75,226
1947	44,771	29,45	592	74,813
1948	46,723	31,53	630	78,883
1949	44,794	31,663	910	77,367
1950	45,754	31,941	1,083	78,778
1951	45,035	32,493	1,152	78,68
1952	46,691	34,156	1,404	82,251
1953	49,569	36,643	1,493	87,705
1954	53,925	40,12	714	94,759
1955	57,005	43,324	696	101,025
1956	59,611	46,745	723	107,079
1957	63,108	50,002	831	113,941
Total	1,281,863	709,035	15,316	2,006,214

Table C.5: Life expectancy at 50 by lifetime income quintile and year of birth with 95% CIs - Men

Year of birth	Quintile 1		Quintile 2		Quintile 3		Quintile 4		Quintile 5	
1930	27.30	[25.40; 30.31]	27.30	[25.47; 29.62]	29.51	[29.08; 29.93]	28.00	[25.90; 30.87]	31.40	[30.92; 31.87]
1931	28.44	[26.23; 31.04]	29.17	[28.78; 29.53]	27.79	[25.75; 30.48]	30.30	[29.80; 30.79]	31.51	[31.02; 31.98]
1932	30.22	[29.97; 30.47]	29.51	[29.22; 29.79]	29.81	[29.34; 30.26]	30.20	[29.75; 30.64]	31.95	[31.61; 32.29]
1933	30.30	[30.16; 30.43]	29.71	[29.56; 29.85]	30.27	[30.08; 30.45]	30.81	[30.61; 31.00]	32.37	[32.16; 32.57]
1934	30.37	[30.16; 30.58]	30.13	[29.98; 30.29]	30.68	[30.46; 30.89]	31.33	[31.17; 31.48]	32.84	[32.63; 33.04]
1935	30.55	[30.34; 30.76]	30.59	[30.38; 30.79]	30.91	[30.69; 31.13]	32.05	[31.81; 32.28]	33.25	[33.03; 33.47]
1936	30.40	[30.18; 30.62]	30.70	[30.47; 30.93]	31.25	[30.98; 31.51]	32.08	[31.87; 32.28]	33.74	[33.58; 33.89]
1937	30.51	[30.29; 30.73]	31.04	[30.77; 31.30]	31.74	[31.51; 31.96]	32.51	[32.26; 32.76]	34.04	[33.80; 34.27]
1938	30.49	[30.27; 30.70]	31.43	[31.16; 31.70]	31.83	[31.58; 32.06]	32.88	[32.59; 33.16]	34.33	[34.04; 34.61]
1939	30.72	[30.49; 30.95]	31.56	[31.33; 31.79]	32.27	[32.04; 32.50]	32.90	[32.54; 33.23]	34.40	[34.13; 34.66]
1940	30.83	[30.63; 31.04]	31.57	[31.30; 31.83]	32.20	[31.92; 32.46]	33.14	[32.85; 33.42]	34.72	[34.48; 34.95]
1941	30.62	[30.36; 30.87]	31.58	[31.33; 31.82]	32.41	[32.08; 32.72]	33.23	[32.86; 33.59]	35.00	[34.70; 35.27]
1942	30.61	[30.36; 30.85]	31.52	[31.24; 31.79]	32.62	[32.32; 32.91]	33.17	[32.86; 33.47]	35.03	[34.72; 35.32]
1943	30.30	[30.11; 30.48]	31.32	[31.06; 31.56]	32.68	[32.37; 32.98]	33.27	[32.69; 33.81]	34.92	[34.65; 35.17]
1944	30.40	[30.08; 30.72]	31.99	[31.54; 32.41]	33.01	[32.65; 33.35]	33.87	[33.49; 34.23]	35.11	[34.72; 35.47]
1945	30.73	[30.46; 30.99]	31.98	[31.57; 32.37]	33.21	[32.69; 33.70]	34.05	[33.68; 34.39]	35.33	[34.90; 35.72]
1946	30.83	[30.52; 31.14]	32.25	[31.91; 32.58]	32.84	[32.41; 33.25]	34.47	[34.08; 34.85]	35.61	[35.34; 35.88]
1947	30.97	[30.54; 31.38]	32.55	[32.08; 33.00]	33.26	[32.77; 33.74]	34.54	[34.00; 35.04]	35.89	[35.36; 36.38]
1948	31.19	[30.87; 31.50]	32.66	[32.26; 33.04]	33.48	[32.58; 34.32]	34.85	[34.27; 35.38]	36.65	[36.31; 36.96]
1949	31.37	[30.85; 31.88]	32.79	[32.21; 33.34]	34.15	[33.53; 34.74]	35.02	[34.42; 35.57]	36.55	[36.01; 37.05]
1950	31.46	[31.04; 31.87]	32.28	[31.51; 33.03]	33.31	[32.69; 33.92]	35.07	[34.53; 35.58]	36.92	[36.40; 37.40]
1951	31.36	[31.03; 31.68]	33.63	[32.93; 34.30]	34.05	[33.29; 34.77]	35.33	[34.53; 36.07]	36.12	[35.27; 36.89]
1952	31.32	[30.80; 31.84]	33.37	[32.73; 33.98]	33.74	[33.09; 34.37]	34.89	[34.11; 35.62]	37.14	[36.26; 37.91]
1953	31.51	[30.76; 32.23]	33.15	[32.53; 33.74]	33.67	[32.76; 34.53]	35.23	[34.26; 36.11]	37.38	[36.63; 38.06]
1954	31.04	[29.93; 32.12]	31.52	[30.17; 32.84]	33.73	[32.48; 34.88]	35.05	[33.92; 36.08]	36.52	[35.47; 37.44]
1955	31.68	[30.74; 32.58]	33.46	[32.32; 34.51]	32.48	[31.11; 33.78]	35.29	[33.65; 36.69]	36.92	[35.84; 37.86]
1956	31.93	[30.86; 32.95]	33.04	[32.06; 33.97]	33.17	[31.03; 35.11]	35.32	[33.91; 36.56]	35.12	[33.11; 36.88]
1957	33.13	[32.28; 33.92]	31.94	[30.45; 33.37]	32.92	[30.52; 35.14]	33.19	[30.90; 35.34]	37.82	[36.66; 38.79]

Table C.6: Life expectancy at 50 by lifetime income quintile and year of birth with 95% CIs - Women

Year of birth	Quintile 1		Quintile 2		Quintile 3		Quintile 4		Quintile 5	
1930	36.58	[35.77; 37.31]	34.97	[32.60; 36.99]	35.88	[34.70; 36.91]	36.15	[35.35; 36.86]	36.21	[35.63; 36.74]
1931	35.29	[33.53; 36.85]	36.82	[36.26; 37.33]	35.91	[34.91; 36.80]	36.53	[35.88; 37.10]	36.05	[35.04; 36.93]
1932	36.83	[36.25; 37.35]	36.41	[35.87; 36.90]	36.58	[35.51; 37.47]	36.27	[35.69; 36.80]	36.56	[35.80; 37.22]
1933	37.06	[36.62; 37.47]	37.07	[36.72; 37.39]	36.47	[36.01; 36.89]	36.25	[35.58; 36.87]	36.90	[36.27; 37.45]
1934	36.91	[36.49; 37.29]	36.76	[36.41; 37.09]	37.48	[37.07; 37.84]	36.83	[36.27; 37.32]	36.85	[36.24; 37.39]
1935	37.24	[36.95; 37.52]	37.65	[37.32; 37.95]	37.66	[37.28; 38.00]	37.33	[36.73; 37.86]	37.28	[36.88; 37.64]
1936	37.55	[37.17; 37.89]	37.54	[37.16; 37.90]	37.75	[37.28; 38.17]	37.60	[37.26; 37.90]	37.15	[36.71; 37.54]
1937	37.42	[36.96; 37.83]	38.17	[37.87; 38.44]	38.09	[37.83; 38.32]	37.37	[36.81; 37.86]	37.54	[37.18; 37.86]
1938	37.78	[37.38; 38.13]	37.56	[37.25; 37.84]	37.97	[37.61; 38.29]	37.68	[37.23; 38.08]	37.79	[37.49; 38.06]
1939	37.56	[37.21; 37.89]	37.84	[37.36; 38.27]	37.94	[37.58; 38.27]	37.66	[37.04; 38.21]	37.81	[37.39; 38.18]
1940	37.17	[36.72; 37.59]	37.94	[37.60; 38.25]	38.17	[37.77; 38.53]	37.79	[37.31; 38.21]	37.66	[37.09; 38.16]
1941	36.99	[36.50; 37.43]	37.80	[37.15; 38.35]	38.15	[37.69; 38.56]	37.87	[37.40; 38.29]	37.62	[37.19; 38.01]
1942	37.05	[36.62; 37.46]	37.91	[37.63; 38.18]	38.04	[37.48; 38.54]	37.86	[37.36; 38.30]	38.16	[37.70; 38.58]
1943	37.33	[36.84; 37.77]	37.46	[36.70; 38.12]	38.33	[37.82; 38.78]	38.08	[37.59; 38.51]	37.58	[37.08; 38.03]
1944	37.16	[36.31; 37.89]	38.44	[38.08; 38.78]	38.38	[37.86; 38.85]	37.63	[37.15; 38.05]	38.26	[37.82; 38.66]
1945	37.75	[37.12; 38.30]	37.80	[37.19; 38.34]	38.02	[37.61; 38.40]	37.73	[36.89; 38.45]	38.34	[37.72; 38.88]
1946	37.43	[36.68; 38.09]	38.70	[38.16; 39.18]	38.57	[37.89; 39.15]	38.72	[38.40; 39.02]	38.50	[38.12; 38.84]
1947	37.70	[37.28; 38.10]	38.48	[37.80; 39.06]	38.84	[38.24; 39.36]	38.58	[38.00; 39.09]	39.61	[38.93; 40.16]
1948	38.06	[37.47; 38.59]	38.79	[38.26; 39.25]	39.13	[38.25; 39.85]	38.45	[37.79; 39.03]	39.38	[38.82; 39.86]
1949	37.89	[37.00; 38.65]	38.73	[38.17; 39.24]	38.79	[38.08; 39.41]	38.28	[37.32; 39.10]	39.20	[38.15; 40.04]
1950	37.67	[36.85; 38.41]	38.12	[37.22; 38.91]	38.63	[37.58; 39.51]	37.58	[36.75; 38.33]	39.63	[38.75; 40.34]
1951	37.46	[36.59; 38.24]	39.63	[38.56; 40.46]	38.70	[36.96; 40.00]	37.40	[36.17; 38.46]	39.86	[39.03; 40.52]
1952	38.28	[37.54; 38.92]	38.65	[37.55; 39.56]	37.27	[35.53; 38.71]	39.33	[38.08; 40.30]	39.49	[38.63; 40.19]
1953	38.06	[37.05; 38.93]	38.86	[37.59; 39.87]	38.50	[36.89; 39.76]	38.88	[37.90; 39.70]	40.24	[39.31; 40.97]
1954	37.67	[36.39; 38.74]	37.99	[36.30; 39.32]	40.05	[38.95; 40.88]	39.03	[37.51; 40.20]	39.99	[38.58; 41.01]
1955	38.91	[37.91; 39.73]	38.26	[36.51; 39.61]	40.12	[38.33; 41.32]	37.79	[35.77; 39.36]	39.94	[37.30; 41.51]
1956	36.48	[33.92; 38.50]	39.00	[37.79; 39.96]	39.80	[37.38; 41.32]	38.18	[35.78; 39.94]	39.17	[36.97; 40.71]
1957	37.09	[34.53; 39.06]	40.57	[39.68; 41.27]	37.27	[34.59; 39.32]	35.64	[32.54; 38.25]	39.27	[37.09; 40.80]

Table C.7: Lifetable entropy at 50 by lifetime income quintile and year of birth with 95% CIs - Men

YoB	Quintile 1		Quintile 2		Quintile 3		Quintile 4		Quintile 5	
1930	0.2804	[0.2260; 0.3264]	0.3246	[0.2688; 0.3519]	0.3441	[0.3415; 0.3453]	0.2956	[0.2358; 0.3255]	0.3035	[0.3023; 0.3037]
1931	0.3151	[0.2583; 0.3326]	0.3545	[0.3522; 0.3556]	0.3080	[0.2494; 0.3356]	0.3255	[0.3230; 0.3262]	0.3031	[0.3021; 0.3033]
1932	0.3402	[0.3398; 0.3403]	0.3509	[0.3498; 0.3515]	0.3333	[0.3303; 0.3347]	0.3252	[0.3228; 0.3263]	0.2991	[0.2982; 0.2995]
1933	0.3442	[0.3440; 0.3443]	0.3516	[0.3513; 0.3518]	0.3346	[0.3342; 0.3347]	0.3203	[0.3199; 0.3204]	0.2934	[0.2927; 0.2940]
1934	0.3359	[0.3356; 0.3360]	0.3376	[0.3372; 0.3378]	0.3247	[0.3244; 0.3249]	0.3114	[0.3113; 0.3115]	0.2855	[0.2845; 0.2862]
1935	0.3317	[0.3315; 0.3318]	0.3290	[0.3288; 0.3291]	0.3162	[0.3158; 0.3164]	0.3011	[0.3003; 0.3016]	0.2816	[0.2802; 0.2828]
1936	0.3335	[0.3331; 0.3336]	0.3239	[0.3235; 0.3241]	0.3116	[0.3112; 0.3117]	0.2975	[0.2969; 0.2979]	0.2722	[0.2711; 0.2731]
1937	0.3279	[0.3274; 0.3281]	0.3176	[0.3172; 0.3177]	0.3063	[0.3057; 0.3067]	0.2901	[0.2891; 0.2907]	0.2647	[0.2631; 0.2662]
1938	0.3252	[0.3246; 0.3255]	0.3126	[0.3120; 0.3129]	0.3001	[0.2997; 0.3003]	0.2812	[0.2800; 0.2820]	0.2588	[0.2567; 0.2608]
1939	0.325 0	[0.3247; 0.3252]	0.3098	[0.3092; 0.3100]	0.2925	[0.2918; 0.2929]	0.2815	[0.2799; 0.2825]	0.2564	[0.2545; 0.2582]
1940	0.3234	[0.3232; 0.3235]	0.3089	[0.3083; 0.3091]	0.2959	[0.2950; 0.2964]	0.2789	[0.2775; 0.2801]	0.2529	[0.2510; 0.2547]
1941	0.3253	[0.3247; 0.3254]	0.3064	[0.3059; 0.3065]	0.2912	[0.2900; 0.2919]	0.2743	[0.2726; 0.2756]	0.2466	[0.2441; 0.2488]
1942	0.3266	[0.3262; 0.3268]	0.3084	[0.3078; 0.3086]	0.2881	[0.2869; 0.2889]	0.2740	[0.2727; 0.2750]	0.2477	[0.2450; 0.2502]
1943	0.3310	[0.3307; 0.3312]	0.3099	[0.3094; 0.3101]	0.2869	[0.2854; 0.2880]	0.2736	[0.2703; 0.2756]	0.2482	[0.2458; 0.2503]
1944	0.3305	[0.3300; 0.3307]	0.3014	[0.2993; 0.3027]	0.2817	[0.2795; 0.2835]	0.2662	[0.2632; 0.2687]	0.2431	[0.2395; 0.2463]
1945	0.3253	[0.3245; 0.3258]	0.3005	[0.2984; 0.3019]	0.2795	[0.2756; 0.2826]	0.2630	[0.2599; 0.2657]	0.2419	[0.2375; 0.2461]
1946	0.3253	[0.3239; 0.3261]	0.2970	[0.2949; 0.2988]	0.2778	[0.2753; 0.2797]	0.2565	[0.2527; 0.2600]	0.2350	[0.2320; 0.2379]
1947	0.3199	[0.3179; 0.3212]	0.2922	[0.2887; 0.2951]	0.2747	[0.2708; 0.2779]	0.2550	[0.2497; 0.2599]	0.2294	[0.2235; 0.2351]
1948	0.3154	[0.3137; 0.3168]	0.2892	[0.2861; 0.2919]	0.2721	[0.2641; 0.2782]	0.2497	[0.2436; 0.2553]	0.2184	[0.2141; 0.2227]
1949	0.3161	[0.3124; 0.3190]	0.2859	[0.2811; 0.2899]	0.2629	[0.2564; 0.2687]	0.2456	[0.2391; 0.2517]	0.2181	[0.2113; 0.2247]
1950	0.3142	[0.3110; 0.3168]	0.2916	[0.2856; 0.2960]	0.2709	[0.2656; 0.2753]	0.2451	[0.2389; 0.2509]	0.2137	[0.2067; 0.2207]
1951	0.3120	[0.3098; 0.3140]	0.2750	[0.2673; 0.2817]	0.2596	[0.2519; 0.2661]	0.2384	[0.2291; 0.2469]	0.2223	[0.2122; 0.2318]
1952	0.3133	[0.3093; 0.3165]	0.2786	[0.2717; 0.2847]	0.2616	[0.2555; 0.2668]	0.2423	[0.2339; 0.2497]	0.2089	[0.1971; 0.2209]
1953	0.3112	[0.3047; 0.3162]	0.2784	[0.2721; 0.2837]	0.2638	[0.2547; 0.2711]	0.2382	[0.2270; 0.2481]	0.2033	[0.1929; 0.2139]
1954	0.3131	[0.3043; 0.3180]	0.2884	[0.2796; 0.2914]	0.2617	[0.2491; 0.2711]	0.2402	[0.2274; 0.2514]	0.2155	[0.2025; 0.2280]
1955	0.3067	[0.2978; 0.3133]	0.2744	[0.2617; 0.2849]	0.2730	[0.2623; 0.2784]	0.2377	[0.2184; 0.2542]	0.2077	[0.1940; 0.2211]
1956	0.3054	[0.2941; 0.3141]	0.2763	[0.2664; 0.2842]	0.2667	[0.2457; 0.2770]	0.2353	[0.2188; 0.2496]	0.2297	[0.2090; 0.2439]
1957	0.2918	[0.2808; 0.3019]	0.2872	[0.2746; 0.2938]	0.2671	[0.2444; 0.2759]	0.2552	[0.2353; 0.2620]	0.1949	[0.1792; 0.2118]

Table C.8: Lifetable entropy at 50 by lifetime income quintile and year of birth with 95% CIs - Women

Year of birth	Quintile 1	Quintile 2	Quintile 3	Quintile 4	Quintile 5
1930	0.2192 [0.2134; 0.2235]	0.2336 [0.2182; 0.2343]	0.2308 [0.2231; 0.2343]	0.2334 [0.2273; 0.2383]	0.2321 [0.2276; 0.2359]
1931	0.2248 [0.2158; 0.2255]	0.2241 [0.2191; 0.2286]	0.2310 [0.2244; 0.2347]	0.2272 [0.2219; 0.2317]	0.2330 [0.2256; 0.2381]
1932	0.2191 [0.2145; 0.2232]	0.2277 [0.2235; 0.2314]	0.2284 [0.2194; 0.2357]	0.2286 [0.2244; 0.2322]	0.2297 [0.2231; 0.2354]
1933	0.2182 [0.2143; 0.2219]	0.2238 [0.2203; 0.2271]	0.2312 [0.2272; 0.2349]	0.2276 [0.2227; 0.2314]	0.2247 [0.2189; 0.2300]
1934	0.2207 [0.2171; 0.2241]	0.2290 [0.2256; 0.2322]	0.2191 [0.2148; 0.2235]	0.2266 [0.2214; 0.2314]	0.2258 [0.2202; 0.2310]
1935	0.2183 [0.2154; 0.2211]	0.2168 [0.2131; 0.2206]	0.2161 [0.2120; 0.2204]	0.2190 [0.2130; 0.2248]	0.2219 [0.2177; 0.2260]
1936	0.2136 [0.2098; 0.2174]	0.2161 [0.2120; 0.2202]	0.2138 [0.2088; 0.2190]	0.2149 [0.2114; 0.2185]	0.2232 [0.2188; 0.2276]
1937	0.2157 [0.2112; 0.2201]	0.2068 [0.2033; 0.2104]	0.2093 [0.2063; 0.2124]	0.2169 [0.2115; 0.2222]	0.2162 [0.2124; 0.2200]
1938	0.2110 [0.2068; 0.2152]	0.2149 [0.2117; 0.2180]	0.2108 [0.2067; 0.2150]	0.2132 [0.2085; 0.2179]	0.2133 [0.2100; 0.2167]
1939	0.2110 [0.2075; 0.2143]	0.2076 [0.2027; 0.2124]	0.2092 [0.2053; 0.2132]	0.2096 [0.2035; 0.2154]	0.2107 [0.2063; 0.2151]
1940	0.2139 [0.2100; 0.2176]	0.2085 [0.2048; 0.2123]	0.2040 [0.1996; 0.2085]	0.2105 [0.2055; 0.2155]	0.2125 [0.2066; 0.2182]
1941	0.2182 [0.2141; 0.2219]	0.2092 [0.2028; 0.2156]	0.2047 [0.1996; 0.2099]	0.2086 [0.2037; 0.2136]	0.2127 [0.2082; 0.2170]
1942	0.2183 [0.2144; 0.2220]	0.2104 [0.2072; 0.2136]	0.2048 [0.1989; 0.2107]	0.2100 [0.2047; 0.2153]	0.2059 [0.2007; 0.2113]
1943	0.2159 [0.2110; 0.2206]	0.2139 [0.2064; 0.2212]	0.2018 [0.1960; 0.2078]	0.2069 [0.2014; 0.2126]	0.2142 [0.2089; 0.2195]
1944	0.2180 [0.2097; 0.2258]	0.2024 [0.1978; 0.2072]	0.2020 [0.1957; 0.2086]	0.2132 [0.2080; 0.2184]	0.2045 [0.1992; 0.2100]
1945	0.2130 [0.2058; 0.2203]	0.2091 [0.2023; 0.2159]	0.2072 [0.2023; 0.2123]	0.2074 [0.1986; 0.2158]	0.2025 [0.1950; 0.2103]
1946	0.2152 [0.2070; 0.2232]	0.1971 [0.1902; 0.2045]	0.1976 [0.1894; 0.2063]	0.1971 [0.1929; 0.2015]	0.1997 [0.1949; 0.2047]
1947	0.2121 [0.2070; 0.2173]	0.2013 [0.1928; 0.2104]	0.1947 [0.1870; 0.2030]	0.1967 [0.1895; 0.2042]	0.1830 [0.1740; 0.1934]
1948	0.2079 [0.2004; 0.2155]	0.1958 [0.1888; 0.2032]	0.1895 [0.1783; 0.2019]	0.2008 [0.1923; 0.2096]	0.1858 [0.1781; 0.1941]
1949	0.2113 [0.2002; 0.2226]	0.1969 [0.1892; 0.2049]	0.1936 [0.1844; 0.2034]	0.2022 [0.1904; 0.2144]	0.1879 [0.1747; 0.2029]
1950	0.2127 [0.2025; 0.2229]	0.2038 [0.1926; 0.2152]	0.1969 [0.1836; 0.2112]	0.2096 [0.1999; 0.2188]	0.1805 [0.1689; 0.1939]
1951	0.2176 [0.2066; 0.2285]	0.1813 [0.1673; 0.1978]	0.1960 [0.1754; 0.2190]	0.2145 [0.2001; 0.2281]	0.1789 [0.1674; 0.1924]
1952	0.2053 [0.1953; 0.2158]	0.1983 [0.1838; 0.2141]	0.2128 [0.1938; 0.2294]	0.1858 [0.1697; 0.2043]	0.1835 [0.1717; 0.1967]
1953	0.2082 [0.1949; 0.2221]	0.1933 [0.1771; 0.2112]	0.1986 [0.1790; 0.2199]	0.1926 [0.1796; 0.2067]	0.1705 [0.1577; 0.1859]
1954	0.2144 [0.1982; 0.2310]	0.2075 [0.1870; 0.2287]	0.1743 [0.1596; 0.1922]	0.1883 [0.1696; 0.2091]	0.1737 [0.1561; 0.1955]
1955	0.1960 [0.1819; 0.2117]	0.2029 [0.1815; 0.2256]	0.1716 [0.1504; 0.1996]	0.2077 [0.1844; 0.2306]	0.1747 [0.1469; 0.2134]
1956	0.2317 [0.2029; 0.2564]	0.1925 [0.1763; 0.2109]	0.1776 [0.1508; 0.2140]	0.2023 [0.1749; 0.2308]	0.1862 [0.1608; 0.2164]
1957	0.2229 [0.1931; 0.2511]	0.1650 [0.1523; 0.1806]	0.2166 [0.1864; 0.2441]	0.2306 [0.2003; 0.2440]	0.1855 [0.1598; 0.2164]

Table C.9: Life expectancy at 67 by pension income quintile and calendar year with 95% CIs - Men

Calendar Year	Quintile 1		Quintile 2		Quintile 3		Quintile 4		Quintile 5	
1995	14.17	[14.09;14.27]	14.55	[14.44;14.66]	14.23	[14.15;14.32]	14.16	[14.07;14.26]	15.25	[15.13;15.38]
1996	14.29	[14.19;14.38]	14.82	[14.71;14.94]	14.32	[14.23;14.40]	14.20	[14.11;14.30]	15.01	[14.90;15.12]
1997	14.09	[14.01;14.18]	14.61	[14.50;14.72]	14.39	[14.31;14.47]	14.30	[14.20;14.39]	15.01	[14.90;15.12]
1998	14.28	[14.20;14.36]	14.65	[14.55;14.76]	14.44	[14.36;14.52]	14.55	[14.46;14.64]	15.26	[15.16;15.37]
1999	14.51	[14.43;14.60]	14.81	[14.70;14.92]	14.69	[14.60;14.78]	14.71	[14.62;14.80]	15.34	[15.24;15.44]
2000	14.72	[14.63;14.82]	14.87	[14.78;14.97]	14.80	[14.71;14.88]	14.88	[14.80;14.97]	15.59	[15.49;15.68]
2001	14.60	[14.52;14.69]	15.17	[15.07;15.26]	15.19	[15.10;15.27]	15.05	[14.96;15.13]	16.01	[15.92;16.11]
2002	14.45	[14.37;14.54]	14.92	[14.82;15.01]	14.91	[14.82;14.99]	15.00	[14.92;15.09]	15.84	[15.75;15.92]
2003	15.05	[14.96;15.13]	15.35	[15.27;15.44]	15.37	[15.29;15.46]	15.58	[15.50;15.67]	16.60	[16.50;16.69]
2004	14.90	[14.81;14.98]	15.30	[15.22;15.39]	15.28	[15.19;15.36]	15.47	[15.39;15.55]	16.09	[16.01;16.18]
2005	15.42	[15.33;15.51]	15.63	[15.55;15.71]	15.61	[15.53;15.69]	15.88	[15.81;15.96]	16.49	[16.42;16.57]
2006	15.34	[15.26;15.42]	15.54	[15.45;15.62]	15.77	[15.68;15.85]	15.82	[15.74;15.91]	16.69	[16.61;16.77]
2007	15.40	[15.32;15.47]	15.67	[15.59;15.75]	15.66	[15.58;15.75]	15.93	[15.85;16.01]	16.87	[16.79;16.94]
2008	15.49	[15.42;15.57]	15.67	[15.59;15.75]	15.84	[15.76;15.92]	16.07	[15.99;16.15]	16.86	[16.79;16.94]
2009	15.61	[15.53;15.69]	15.84	[15.77;15.93]	15.98	[15.90;16.05]	16.23	[16.16;16.31]	17.03	[16.96;17.11]
2010	15.64	[15.56;15.72]	16.08	[15.99;16.16]	16.08	[16;16.16.00]	16.54	[16.46;16.62]	17.21	[17.13;17.28]
2011	15.42	[15.34;15.50]	15.85	[15.77;15.93]	16.18	[16.10;16.26]	16.54	[16.46;16.62]	17.16	[17.08;17.24]
2012	15.77	[15.68;15.85]	16.13	[16.05;16.22]	16.44	[16.36;16.52]	16.66	[16.58;16.74]	17.43	[17.34;17.50]
2013	15.94	[15.86;16.01]	16.37	[16.28;16.46]	16.56	[16.48;16.64]	16.99	[16.91;17.07]	17.69	[17.61;17.78]
2014	15.76	[15.68;15.84]	16.06	[15.98;16.15]	16.29	[16.21;16.37]	16.89	[16.81;16.97]	17.64	[17.56;17.72]
2015	16.13	[16.05;16.21]	16.43	[16.34;16.51]	16.69	[16.61;16.77]	17.27	[17.18;17.35]	18.08	[17.99;18.15]
2016	15.87	[15.80;15.95]	16.28	[16.20;16.36]	16.68	[16.60;16.76]	17.28	[17.20;17.35]	17.90	[17.82;17.98]
2017	16.38	[16.30;16.46]	16.66	[16.58;16.74]	17.07	[16.99;17.16]	17.53	[17.45;17.60]	18.34	[18.26;18.42]

Table C.10: Life expectancy at 67 by pension income quintile and calendar year with 95% CIs - Women

Calendar Year	Quntile 1		Quintile 2		Quintile 3		Quintile 4		Quintile 5	
1995	18.70	[18,59;18,80]	18.77	[18,64;18,90]	18.61	[18,45;18,78]	18.34	[18,25;18,42]	18.41	[18,31;18,51]
1996	18.66	[18,56;18,76]	18.95	[18,83;19,06]	18.23	[18,08;18,40]	18.34	[18,26;18,43]	18.30	[18,20;18,40]
1997	18,65	[18,55;18,76]	18.73	[18,62;18,86]	18.41	[18,25;18,55]	18.41	[18,32;18,51]	18.33	[18,23;18,44]
1998	18.85	[18,75;18,95]	18.82	[18,72;18,92]	18.38	[18,23;18,52]	18.34	[18,26;18,42]	18.60	[18,50;18,69]
1999	19.00	[18,90;19,09]	19.00	[18,89;19,10]	18.40	[18,26;18,54]	18.55	[18,46;18,63]	18.47	[18,38;18,57]
2000	19.19	[19,09;19,28]	19.18	[19,07;19,29]	18.83	[18,68;18,97]	18.82	[18,75;18,91]	19.00	[18,91;19,10]
2001	19.43	[19,33;19,51]	19.33	[19,23;19,43]	19.18	[19,03;19,31]	19.09	[19,00;19,17]	18.92	[18,83;19,01]
2002	19.00	[18,92;19,09]	19.17	[19,07;19,26]	18.64	[18,51;18,77]	18.66	[18,58;18,74]	18.77	[18,68;18,86]
2003	19.77	[19,69;19,86]	20.03	[19,93;20,12]	19.42	[19,31;19,55]	19.40	[19,32;19,48]	19.89	[19,80;19,98]
2004	19.58	[19,50;19,66]	19.58	[19,50;19,67]	18.98	[18,87;19,08]	18.99	[18,92;19,07]	19.16	[19,08;19,25]
2005	19.82	[19,73;19,90]	20.17	[20,08;20,25]	19.54	[19,43;19,64]	19.57	[19,49;19,64]	19.42	[19,34;19,50]
2006	20.00	[19,91;20,08]	19.85	[19,77;19,94]	19.59	[19,49;19,69]	19.56	[19,48;19,64]	19.58	[19,50;19,66]
2007	19.96	[19,88;20,04]	19.97	[19,89;20,05]	19.53	[19,43;19,62]	19.55	[19,48;19,64]	19.43	[19,35;19,51]
2008	20.21	[20,13;20,28]	20.03	[19,95;20,11]	19.93	[19,83;20,02]	19.43	[19,36;19,51]	19.48	[19,40;19,56]
2009	20.13	[20,06;20,22]	20.47	[20,38;20,55]	20.01	[19,92;20,09]	19.75	[19,68;19,83]	20.02	[19,94;20,10]
2010	20.07	[19,98;20,15]	20.39	[20,30;20,47]	19.86	[19,78;19,94]	19.94	[19,85;20,01]	20.08	[20,00;20,16]
2011	20.34	[20,26;20,42]	20.13	[20,05;20,21]	19.95	[19,87;20,03]	19.62	[19,54;19,70]	19.85	[19,78;19,94]
2012	20.72	[20,64;20,80]	20.75	[20,68;20,83]	20.11	[20,03;20,20]	20.03	[19,95;20,11]	20.29	[20,21;20,37]
2013	20.84	[20,75;20,91]	20.81	[20,73;20,90]	20.27	[20,19;20,35]	20.37	[20,30;20,45]	20.19	[20,11;20,27]
2014	20.35	[20,26;20,43]	20.41	[20,33;20,49]	19.84	[19,76;19,91]	19.73	[19,64;19,81]	19.81	[19,73;19,88]
2015	20.66	[20,58;20,74]	20.68	[20,60;20,76]	20.14	[20,06;20,22]	20.24	[20,16;20,33]	20.26	[20,19;20,34]
2016	20.54	[20,46;20,61]	20.57	[20,49;20,65]	20.15	[20,07;20,23]	20.06	[19,97;20,13]	20.10	[20,02;20,18]
2017	20.93	[20,84;21,01]	20.80	[20,71;20,88]	20.45	[20,37;20,53]	20.26	[20,18;20,34]	20.42	[20,34;20,50]

Table C.11: Lifetable entropy at 67 by pension income quintile and calendar year with 95% CIs - Men

Calendar Year	Quintile 1	Quintile 2	Quintile 3	Quintile 4	Quintile 5
1995	0.5146 [0.5102;0.5186]	0.4856 [0.4806;0.4902]	0.4989 [0.4946;0.5030]	0.5069 [0.5026;0.5115]	0.5094 [0.5044;0.514]
1996	0.5099 [0.5055;0.5138]	0.4934 [0.4887;0.4978]	0.4978 [0.4937;0.5021]	0.5074 [0.5029;0.5115]	0.4838 [0.4795;0.4879]
1997	0.5125 [0.5081;0.5164]	0.4816 [0.4774;0.4854]	0.4951 [0.4913;0.4992]	0.5067 [0.5027;0.5112]	0.4710 [0.4667;0.4751]
1998	0.4921 [0.4882;0.4962]	0.4889 [0.4849;0.4929]	0.4913 [0.4874;0.4952]	0.5045 [0.5001;0.5083]	0.4789 [0.4748;0.4830]
1999	0.5009 [0.4969;0.5049]	0.4763 [0.4723;0.4804]	0.4821 [0.4781;0.4863]	0.4897 [0.4857;0.4936]	0.4612 [0.4574;0.4651]
2000	0.5081 [0.5038;0.5122]	0.4748 [0.4710;0.4786]	0.4816 [0.4778;0.4853]	0.4774 [0.4738;0.4809]	0.4545 [0.4508;0.4580]
2001	0.4999 [0.4959;0.5037]	0.4761 [0.4724;0.4800]	0.4907 [0.4870;0.4944]	0.4810 [0.4771;0.4849]	0.4619 [0.4585;0.4657]
2002	0.4869 [0.4833;0.4906]	0.4778 [0.4740;0.4815]	0.4737 [0.4700;0.4775]	0.4729 [0.4694;0.4764]	0.4535 [0.4498;0.4570]
2003	0.4868 [0.4831;0.4907]	0.4666 [0.4631;0.4703]	0.4710 [0.4672;0.4747]	0.4680 [0.4645;0.4719]	0.4559 [0.4523;0.4592]
2004	0.4849 [0.4813;0.4889]	0.4614 [0.4576;0.4648]	0.4527 [0.4491;0.4564]	0.4557 [0.4521;0.4592]	0.4261 [0.4225;0.4297]
2005	0.4866 [0.4830;0.4904]	0.4681 [0.4645;0.4719]	0.4560 [0.4525;0.4596]	0.4480 [0.4445;0.4515]	0.4195 [0.4160;0.4229]
2006	0.4744 [0.4703;0.4778]	0.4615 [0.4578;0.4653]	0.4519 [0.4482;0.4558]	0.4347 [0.4311;0.4383]	0.4124 [0.4088;0.4156]
2007	0.4717 [0.4682;0.4755]	0.4549 [0.4512;0.4586]	0.4379 [0.4342;0.4417]	0.4381 [0.4346;0.4415]	0.4080 [0.4046;0.4114]
2008	0.4684 [0.4645;0.4721]	0.4553 [0.4517;0.4590]	0.4453 [0.4413;0.4488]	0.4233 [0.4198;0.4268]	0.4057 [0.4022;0.4092]
2009	0.4545 [0.4507;0.4584]	0.4536 [0.4497;0.4576]	0.4282 [0.4247;0.4316]	0.4180 [0.4144;0.4216]	0.3944 [0.3910;0.3974]
2010	0.4620 [0.4580;0.4655]	0.4542 [0.4505;0.4582]	0.4286 [0.4252;0.4323]	0.4195 [0.4162;0.4231]	0.3906 [0.3873;0.3938]
2011	0.4547 [0.4511;0.4585]	0.4464 [0.4426;0.4499]	0.4286 [0.4250;0.4321]	0.4162 [0.4124;0.4198]	0.3942 [0.3908;0.3977]
2012	0.4491 [0.4452;0.4527]	0.4434 [0.4395;0.4473]	0.4288 [0.4251;0.4323]	0.4066 [0.4031;0.4103]	0.3914 [0.3881;0.3951]
2013	0.4414 [0.4379;0.4449]	0.4278 [0.4238;0.4314]	0.421 [0.4176;0.4245]	0.4081 [0.4044;0.4117]	0.3863 [0.3826;0.3897]
2014	0.4386 [0.4347;0.4424]	0.4377 [0.4338;0.4415]	0.4214 [0.4179;0.4251]	0.4028 [0.3992;0.4063]	0.3877 [0.3843;0.3912]
2015	0.4427 [0.4392;0.4463]	0.4268 [0.4230;0.4305]	0.4210 [0.4173;0.4245]	0.4055 [0.4018;0.4092]	0.3855 [0.3823;0.3889]
2016	0.4365 [0.4327;0.4403]	0.4272 [0.4237;0.4307]	0.4106 [0.4072;0.4140]	0.4058 [0.4025;0.4091]	0.3760 [0.3728;0.3794]
2017	0.4382 [0.4347;0.4417]	0.4279 [0.4242;0.4312]	0.4193 [0.4159;0.4228]	0.3976 [0.3944;0.4008]	0.3798 [0.3765;0.3830]

Table C.12: Lifetable entropy at 67 by pension income quintile and calendar year with 95% CIs - Women

Calendar Year	Quintile 1	Quintile 2	Quintile 3	Quintile 4	Quintile 5
1995	0.3812 [0.3771; 0.3852]	0.3700 [0.3653; 0.3742]	0.3828 [0.3764; 0.3887]	0.3786 [0.3753; 0.3821]	0.3759 [0.3722; 0.3800]
1996	0.3748 [0.3711; 0.3787]	0.3742 [0.3700; 0.3783]	0.3899 [0.3836; 0.3956]	0.3825 [0.3790; 0.3860]	0.3803 [0.3761; 0.3843]
1997	0.3632 [0.3594; 0.3672]	0.3680 [0.3636; 0.3721]	0.3803 [0.3749; 0.3860]	0.3809 [0.3772; 0.3844]	0.3756 [0.3717; 0.3793]
1998	0.3657 [0.3619; 0.3695]	0.3555 [0.3518; 0.3592]	0.3650 [0.3599; 0.3699]	0.3651 [0.3618; 0.3683]	0.3743 [0.3707; 0.3780]
1999	0.363 [0.3595; 0.3665]	0.3538 [0.3502; 0.3577]	0.3613 [0.3564; 0.3661]	0.3592 [0.3560; 0.3626]	0.3609 [0.3574; 0.3646]
2000	0.3605 [0.3571; 0.3639]	0.3560 [0.3521; 0.3596]	0.3713 [0.3665; 0.3759]	0.3634 [0.3600; 0.3665]	0.3679 [0.3645; 0.3715]
2001	0.3601 [0.3567; 0.3637]	0.3570 [0.3535; 0.3605]	0.3690 [0.3648; 0.3736]	0.3603 [0.3570; 0.3634]	0.3635 [0.3600; 0.3667]
2002	0.3526 [0.3492; 0.3557]	0.3485 [0.3453; 0.3516]	0.3609 [0.3567; 0.3651]	0.3481 [0.3448; 0.3513]	0.3665 [0.3631; 0.3698]
2003	0.3533 [0.3499; 0.3566]	0.3444 [0.3412; 0.3480]	0.3557 [0.3517; 0.3594]	0.3496 [0.3466; 0.3529]	0.3687 [0.3652; 0.3720]
2004	0.3512 [0.348; 0.3544]	0.3340 [0.3309; 0.3370]	0.3481 [0.3444; 0.3518]	0.3358 [0.3328; 0.3387]	0.3514 [0.3481; 0.3548]
2005	0.3389 [0.3358; 0.3419]	0.3376 [0.3344; 0.3407]	0.3474 [0.3439; 0.3510]	0.3457 [0.3425; 0.3489]	0.3469 [0.3438; 0.3502]
2006	0.3459 [0.3427; 0.3492]	0.3388 [0.3357; 0.3420]	0.3475 [0.3440; 0.3512]	0.3441 [0.3409; 0.3472]	0.3508 [0.3476; 0.3537]
2007	0.3422 [0.3393; 0.3452]	0.3388 [0.3358; 0.3418]	0.3482 [0.3449; 0.3516]	0.3496 [0.3462; 0.3526]	0.3396 [0.3366; 0.3427]
2008	0.3398 [0.337; 0.3428]	0.3378 [0.3346; 0.3408]	0.3463 [0.3431; 0.3496]	0.3376 [0.3345; 0.3405]	0.3408 [0.3376; 0.3438]
2009	0.3318 [0.3286; 0.3347]	0.3418 [0.3388; 0.3449]	0.3491 [0.3459; 0.3522]	0.3359 [0.3330; 0.3389]	0.3348 [0.3317; 0.3378]
2010	0.3348 [0.3317; 0.3379]	0.3346 [0.3316; 0.3378]	0.3383 [0.3353; 0.3414]	0.3449 [0.3421; 0.3482]	0.3377 [0.3346; 0.3408]
2011	0.3416 [0.3385; 0.3446]	0.3242 [0.3212; 0.3272]	0.3384 [0.3352; 0.3414]	0.3411 [0.3379; 0.3443]	0.3325 [0.3293; 0.3356]
2012	0.3307 [0.3278; 0.3338]	0.3373 [0.3341; 0.3403]	0.3380 [0.3348; 0.3413]	0.3374 [0.3343; 0.3408]	0.3349 [0.3317; 0.3382]
2013	0.3279 [0.3249; 0.3312]	0.3409 [0.3378; 0.3441]	0.3367 [0.3337; 0.3399]	0.3439 [0.3410; 0.3469]	0.3328 [0.3296; 0.3359]
2014	0.3293 [0.3261; 0.3325]	0.3349 [0.3318; 0.3382]	0.3336 [0.3305; 0.3366]	0.3386 [0.3354; 0.3419]	0.3321 [0.3289; 0.3353]
2015	0.3282 [0.325; 0.3312]	0.3317 [0.3288; 0.3349]	0.3348 [0.3313; 0.3380]	0.3363 [0.3329; 0.3395]	0.3327 [0.3296; 0.3358]
2016	0.3297 [0.3267; 0.3329]	0.3418 [0.3388; 0.3451]	0.3416 [0.3385; 0.3450]	0.3369 [0.3340; 0.3401]	0.3337 [0.3306; 0.3369]
2017	0.3224 [0.3193; 0.3257]	0.3371 [0.3337; 0.3402]	0.3293 [0.3262; 0.3326]	0.3373 [0.3343; 0.3403]	0.3337 [0.3306; 0.3369]

Table C.13: Men's post-retirement mortality and wives' pension

	(1) Odds ratio
Men's pension quintile [ref: 1st, bottom]	.
2nd	0.955***
	(0.00688)
3rd	0.932***
	(0.00694)
4th	0.854***
	(0.00661)
5h (top)	0.720***
	(0.00600)
Wife's pension quintile [ref: 1st, bottom]	
2nd	0.985**
	(0.00708)
3rd	0.954***
	(0.00703)
4th	0.973***
	(0.00746)
5th (top)	1.027***
	(0.00857)
Constant	4.73e-05***
	(5.20e-05)
Observations	7,412,143

Notes. Results from logistic survival analysis based on male retirees from the FPLD fund, who retired between 1995 and 2017, whose wife was alive in 1995 and also retired between 1995 and 2017. The mortality follow-up extends from the year men turn 67 to the end of 2018 or the year of their death, if the latter occurs earlier. Dependent variable is a dummy taking value 1 if the woman dies by the end of the year, 0 otherwise. Pension quintiles are cohort-specific for both men and wives. Control variables: year of birth, age difference with respect to wife, widowhood status, macro-region of residence, macro-region of birth, and 23 duration dummies.

Table C.14: Monthly pension minimum (euro), Italy, 1995-2017

Year	Minimum	Year	Minimum
1995	382.4	2007	436.1
1996	382.4	2008	443.6
1997	382.4	2009	457.8
1998	382.4	2010	461.0
1999	382.4	2011	468.4
2000	382.4	2012	481.0
2001	382.4	2013	495.4
2002	392.7	2014	500.9
2003	402.1	2015	501.9
2004	412.2	2016	501.9
2005	420.4	2017	501.9
2006	427.6		

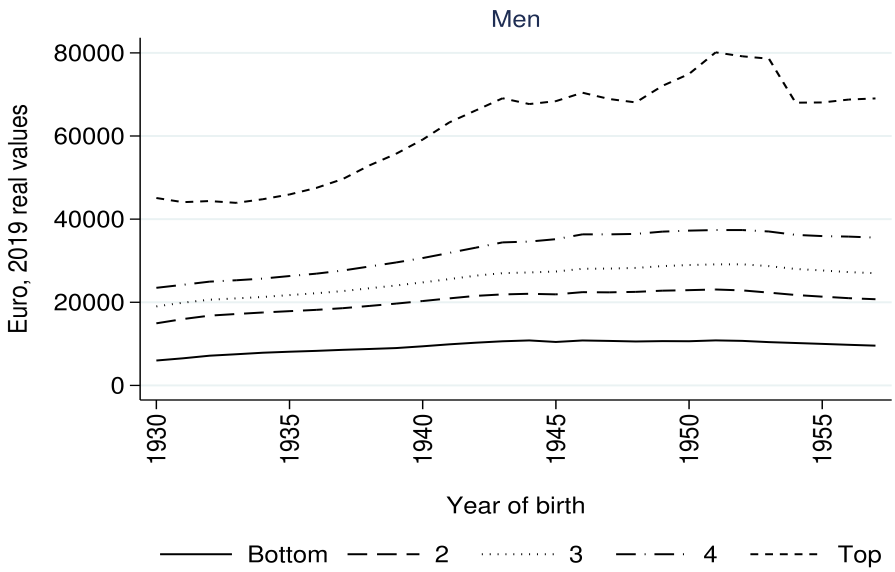
Notes. Source: INPS.

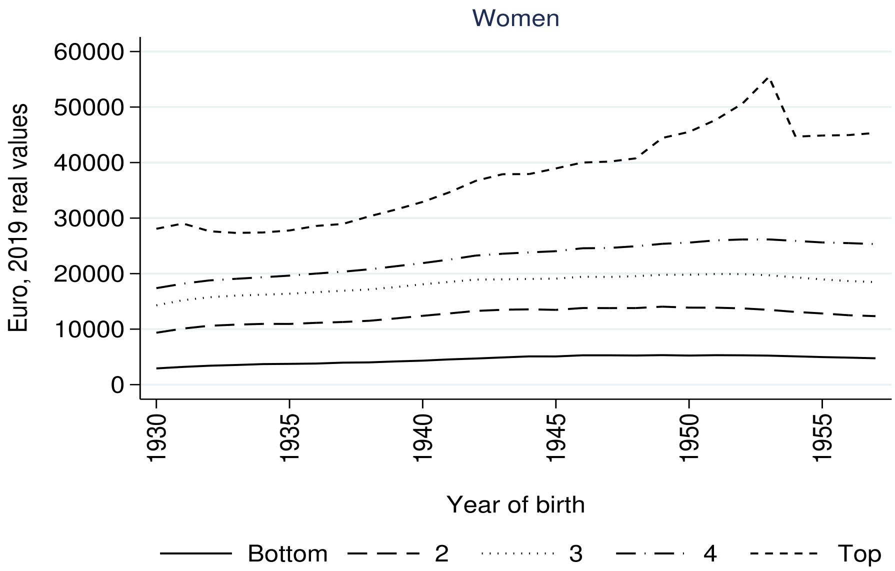
Supplementary Figures

Figure C.1: Permanent earnings by sex and birth cohort

Notes: The graph plots the evolution of inflation-adjusted average private employment earnings between ages 45-49 by year of birth and sex. Earnings are expressed in 2019 euros. Own elaboration based on INPS data.

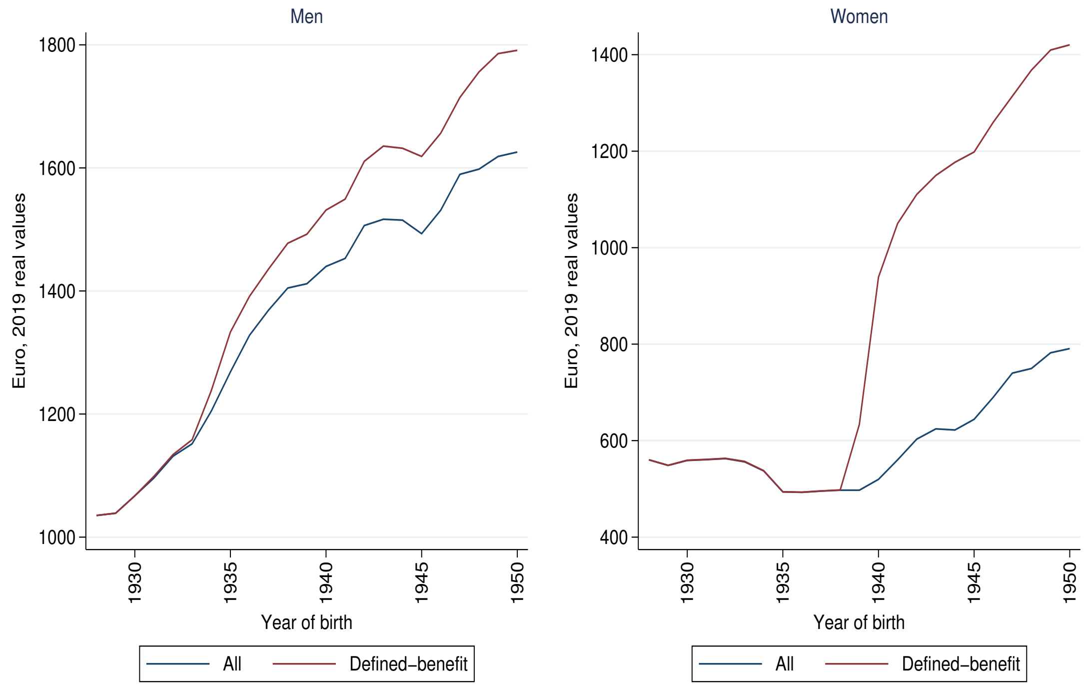
Figure C.2: Permanent earnings by sex, birth cohort and quintile





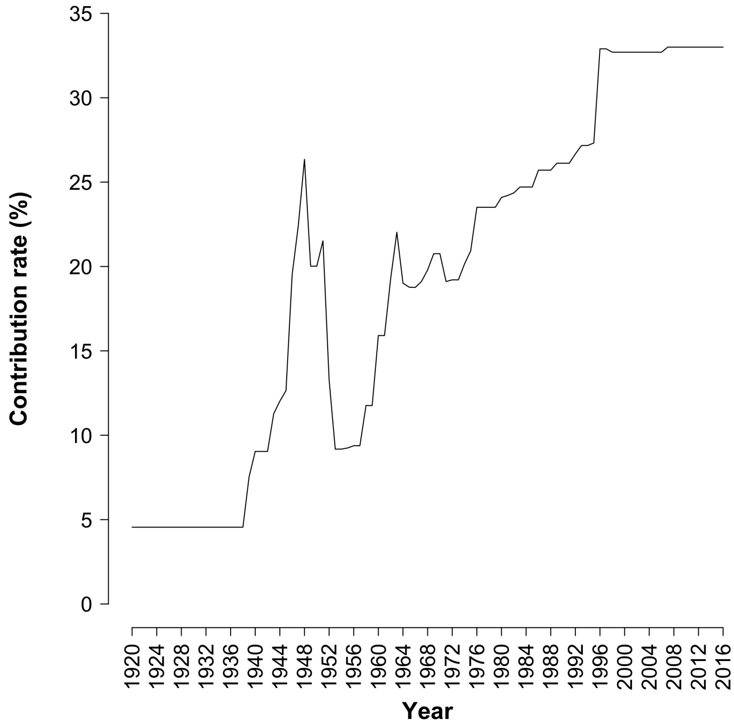
Notes: The graphs plot the evolution of inflation-adjusted average private employment earnings between ages 45-49 by year of birth and quintile. Earnings are expressed in 2019 euros. Own elaboration based on INPS data.

Figure C.3: Gross pension income at age 67 by sex and birth cohort
FPLD pensioners



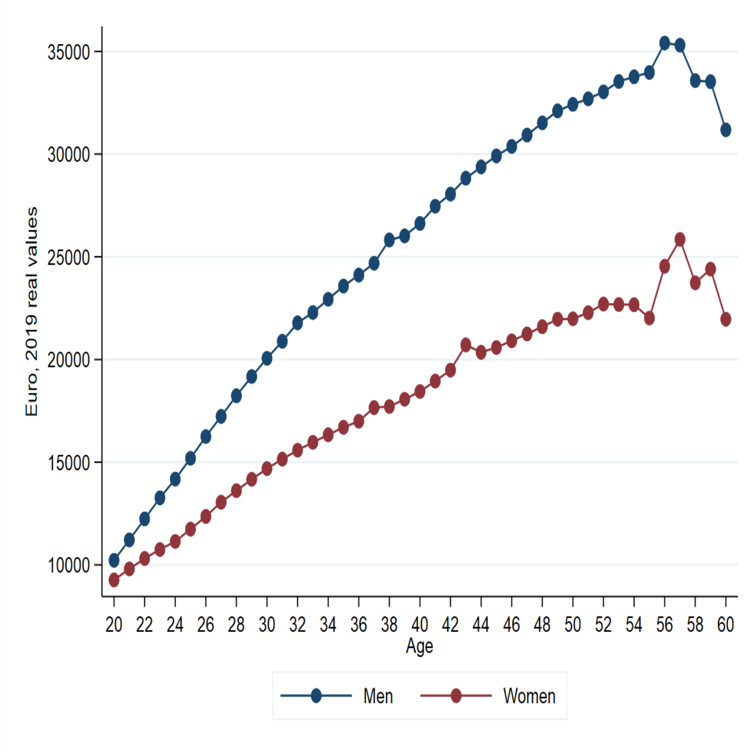
Notes: The graphs plot the evolution of inflation-adjusted median gross pension income at age 67 for pensioners from the private employees pension fund (*Fondo Pensione Lavoratori Dipendenti*, FPLD). Income is expressed in 2019 euros. Own elaboration based on INPS data.

Figure C.4: Contribution rate to *Fondo Pensione Lavoratori Dipendenti* (FPLD fund)



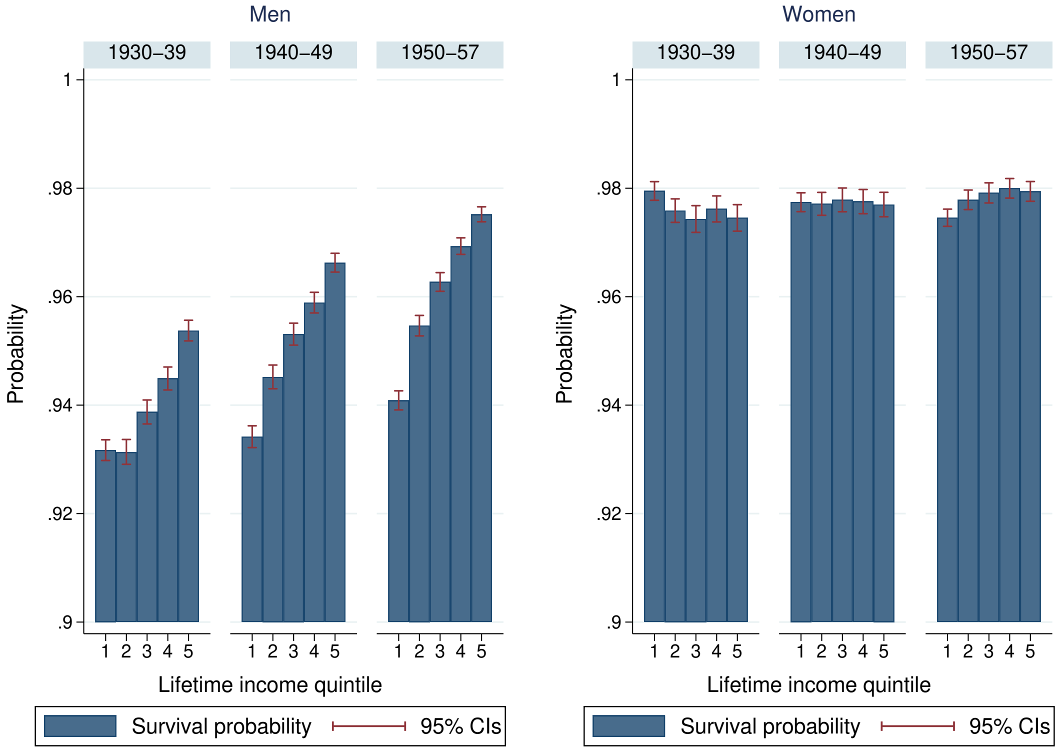
Notes: The graph plots the evolution of contribution rate to the INPS-managed pension fund of private employees. Source: INPS.

Figure C.5: Life cycle earnings profile



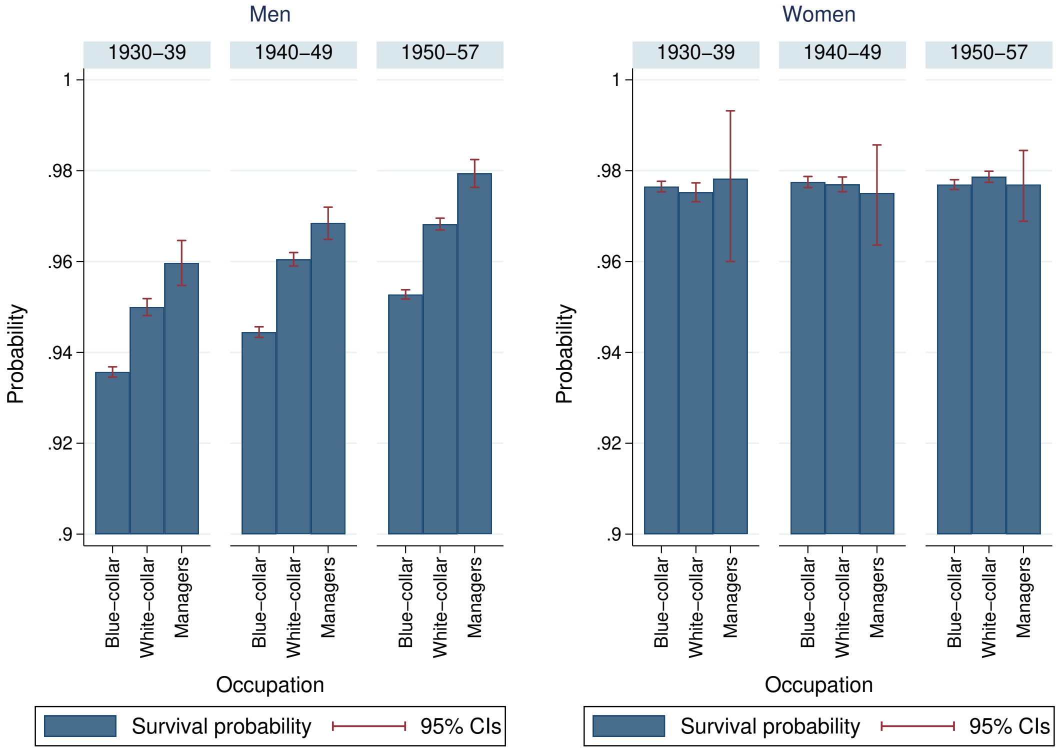
Notes: The graph plots the life cycle earnings profile for individuals employed in the private sector in Italy (birth cohorts 1930-1950). Own elaboration based on INPS data.

Figure C.6: Probability of surviving to 61 at 50 by quintile of average mid-career earnings



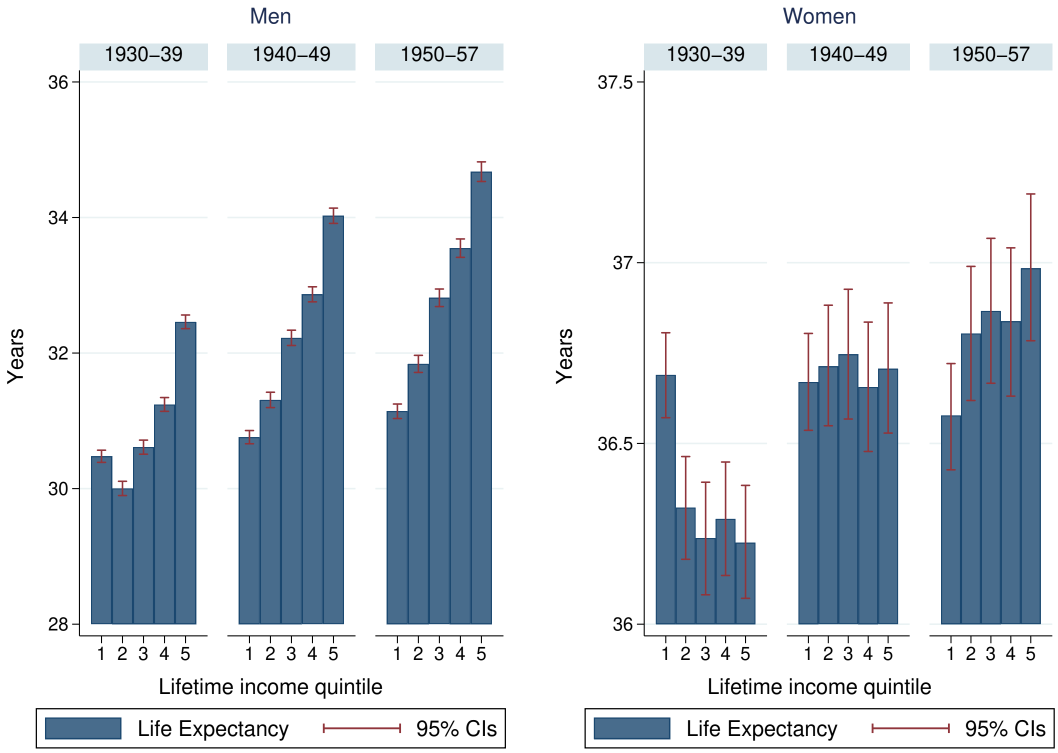
Notes: The graph plots the evolution of survival probability to 61 at 50 by quintile of average employment earnings at 45-49 and birth cohort, along with 95% confidence intervals. Own elaboration based on INPS data.

Figure C.7: Probability of surviving to 61 at 50 by mid-career occupational position



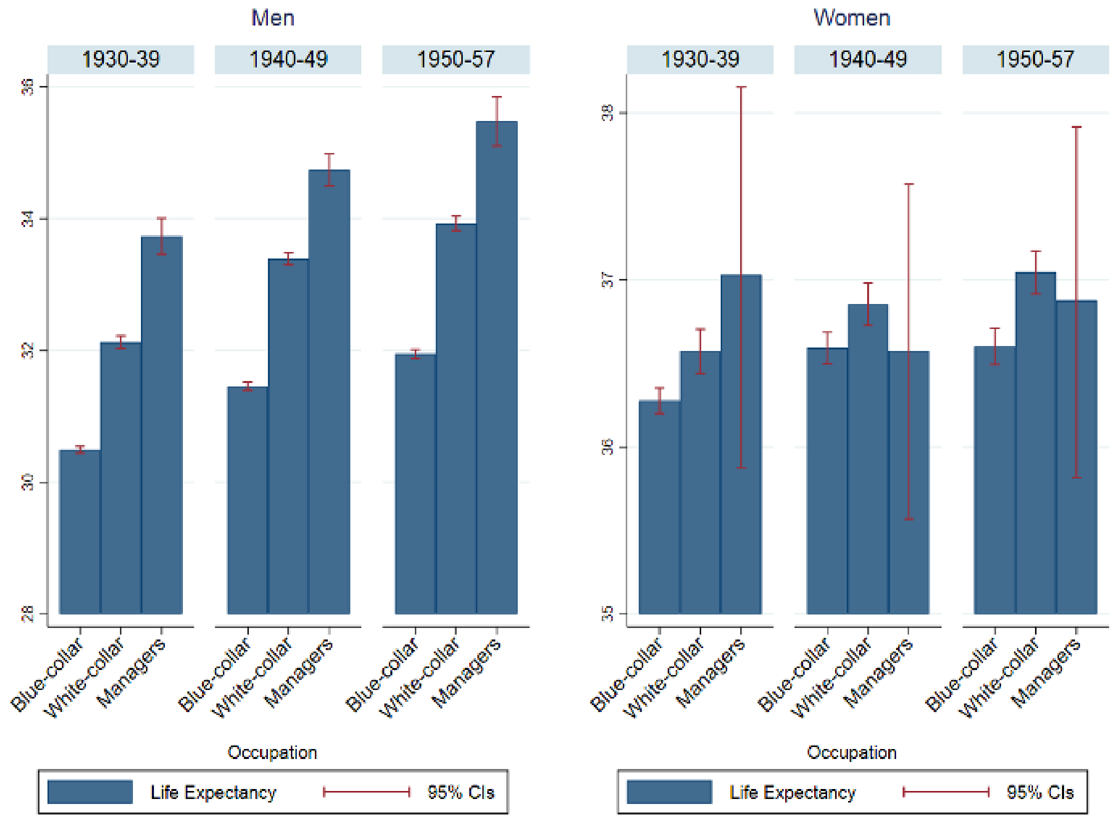
Notes: The graph plots the evolution of survival probability to 61 at 50 by quintile of average employment earnings at 45-49 and birth cohort, along with 95% confidence intervals. Own elaboration based on INPS data.

Figure C.8: Life expectancy at age 50 by quintile of average mid-career earnings based on observed mortality rates



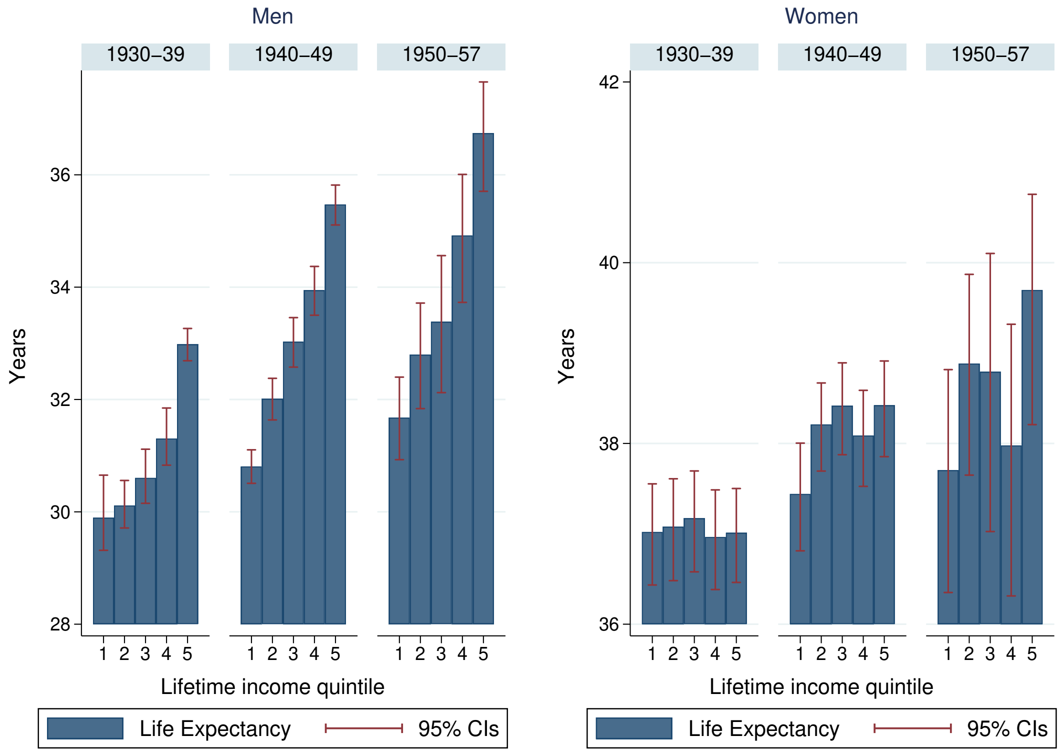
Notes: The graph plots the evolution of life expectancy at 50 by quintile of average employment earnings at 45-49 and birth cohort, along with 95% confidence intervals. Life expectancy estimates are based on observed survival probabilities only. For later cohorts, survival probabilities of earlier cohorts were employed (see Section 3.3.4.2). Own elaboration based on INPS data.

Figure C.9: Life expectancy at age 50 by mid-career occupational position based on observed mortality rates



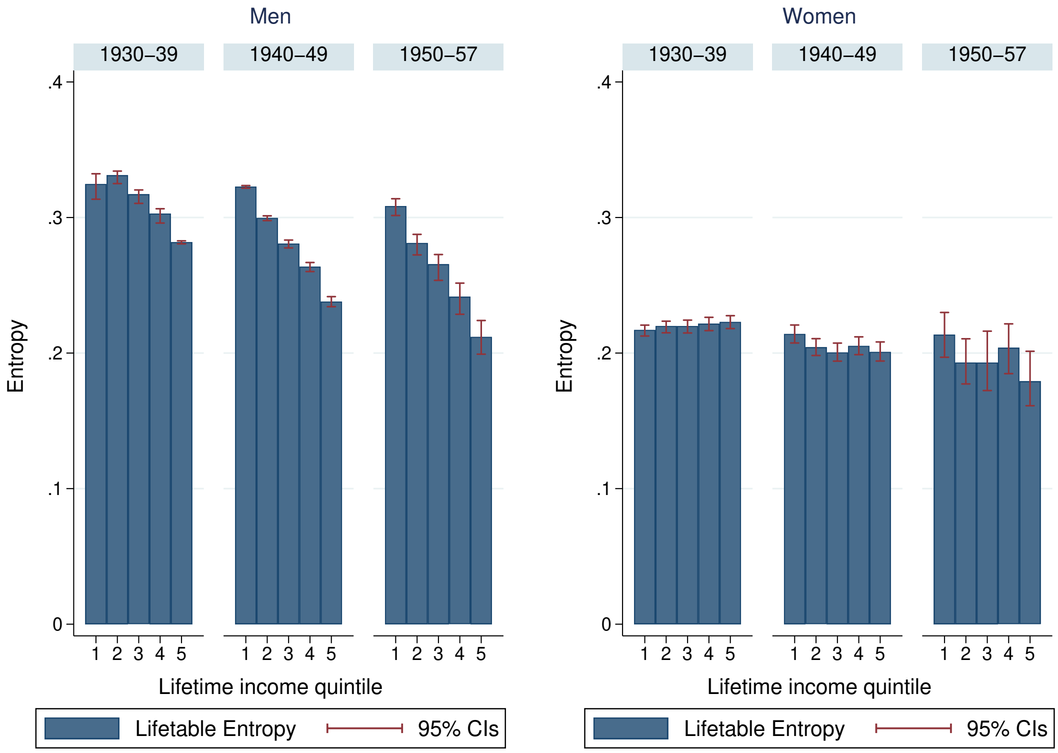
Notes: The graph plots the evolution of life expectancy at 50 by by mid-career occupational position (private employees) and birth cohort, along with 95% confidence intervals. Life expectancy estimates are based on observed survival probabilities only. For later cohorts, survival probabilities of earlier cohorts were employed (see Section **??**). Own elaboration based on INPS data.

Figure C.10: Life expectancy at age 50 by average mid-career earnings based on individuals observed for at least 4 years between ages 45-49



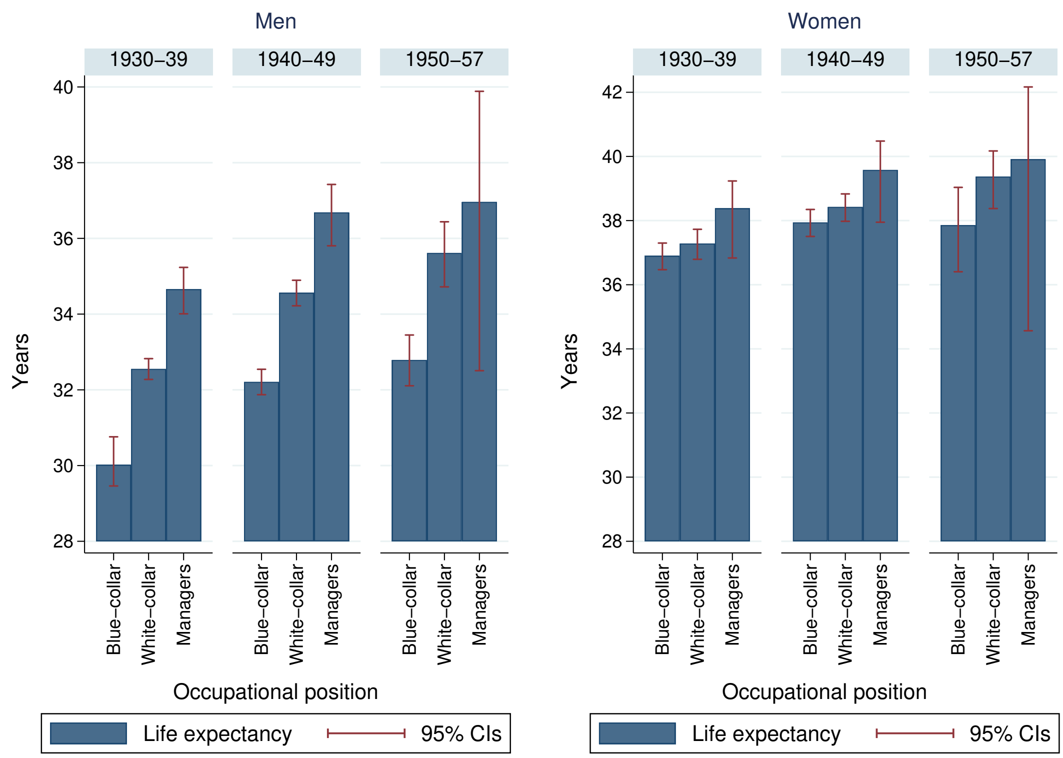
Notes: The graph plots the evolution of life expectancy at 50 by average mid-career private employment earnings and birth cohort, along with 95% confidence intervals. Estimates are constructed starting from a sample which includes only individuals who are observed for at least 4 years between ages 45-49 in the *Dichiarazioni Uniemens* archive. Own elaboration based on INPS data.

Figure C.11: Lifetable entropy at age 50 by average mid-career earnings based on individuals observed for at least 4 years between ages 45-49



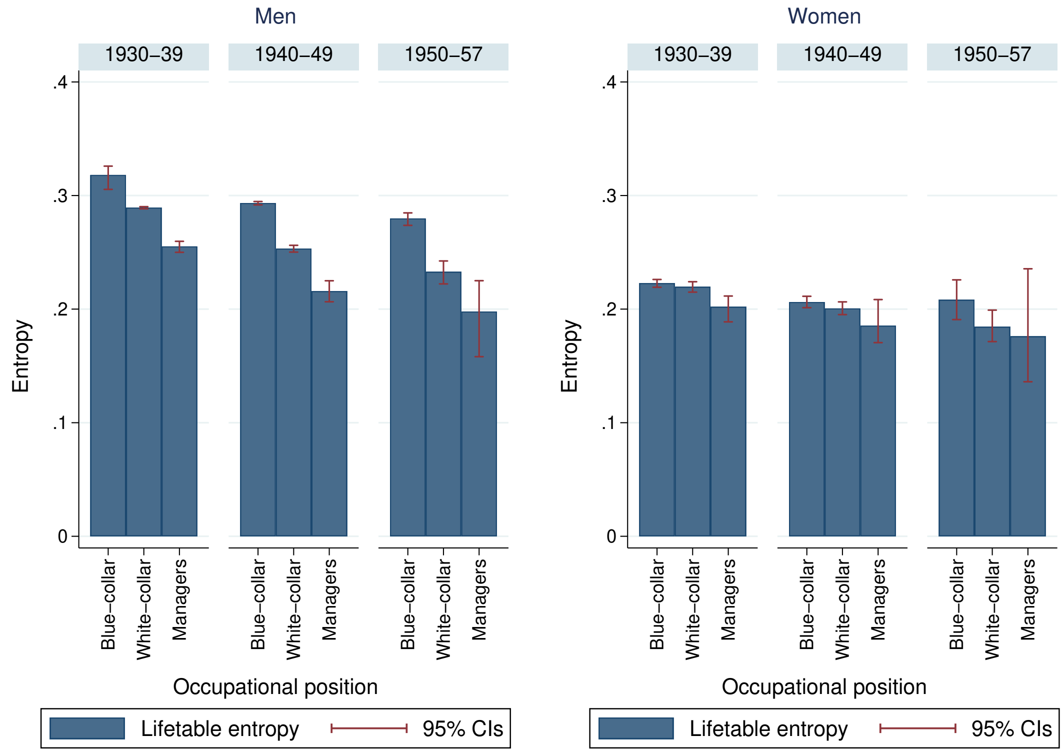
Notes: The graph plots the evolution of lifetable entropy at 50 by average mid-career private employment earnings and birth cohort, along with 95% confidence intervals. Estimates are constructed starting from a sample which includes only individuals who are observed for at least 4 years between ages 45-49 in the *Dichiarazioni Uniemens* archive. Own elaboration based on INPS data.

Figure C.12: Life expectancy at age 50 by mid-career occupation based on individuals observed for at least 4 years between ages 45-49



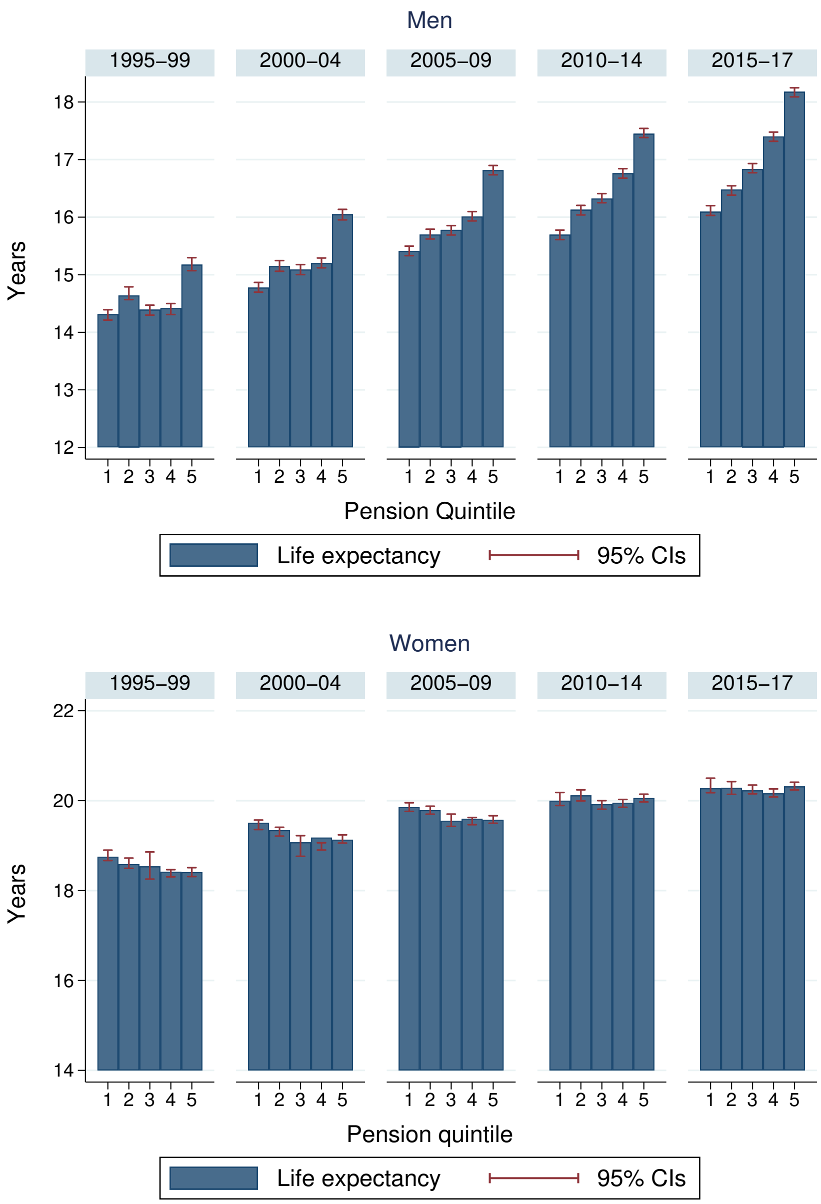
Notes: The graph plots the evolution of life expectancy at 50 by mid-career occupation (private employees), sex and birth cohort, along with 95% confidence intervals. Estimates are constructed starting from a sample which includes only individuals who are observed for at least 4 years between ages 45-49 in the *Dichiarazioni Uniemens* archive. Own elaboration based on INPS data.

Figure C.13: Lifetable entropy at age 50 by mid-career occupation based on individuals observed for at least 4 years between ages 45-49



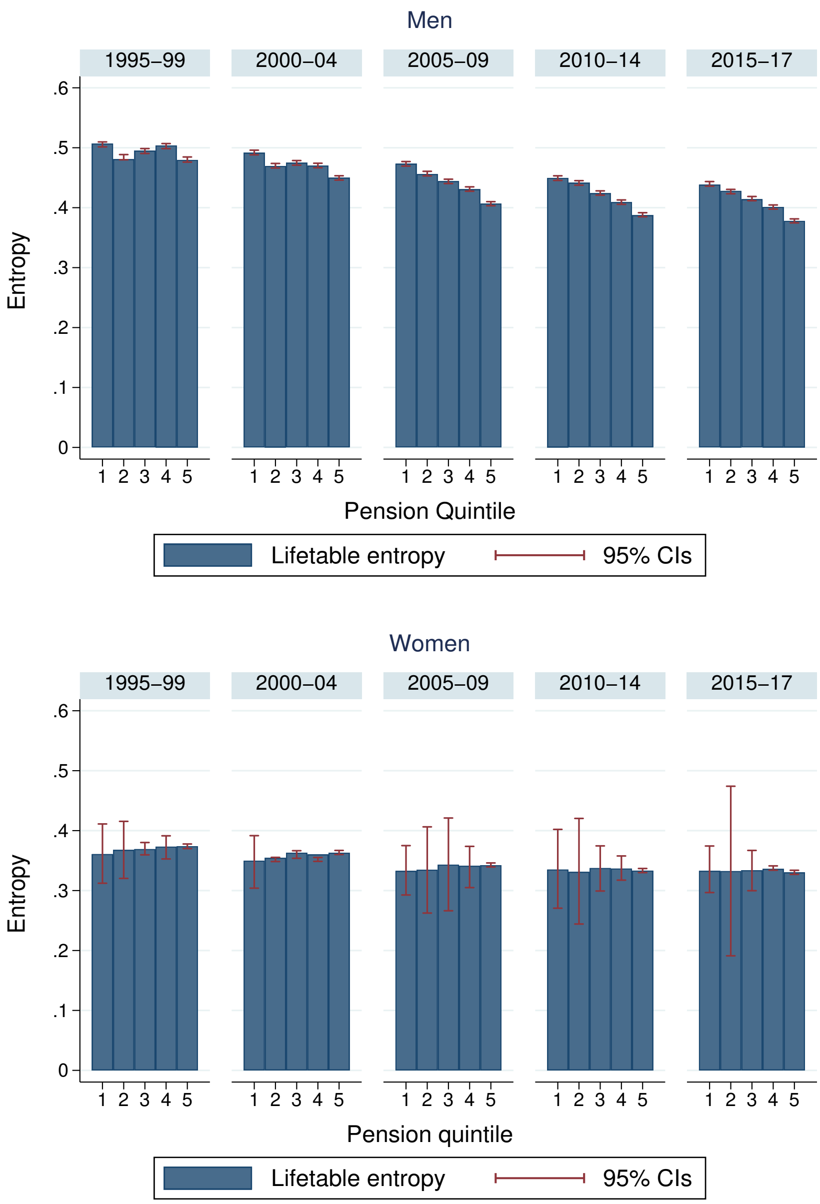
Notes: The graph plots the evolution of lifetable entropy at 50 by mid-career occupation (private employees), sex and birth cohort, along with 95% confidence intervals. Estimates are constructed starting from a sample which includes only individuals who are observed for at least 4 years between ages 45-49 in the *Dichiarazioni Uniemens* archive. Own elaboration based on INPS data.

Figure C.14: Life expectancy by pension income at 67 - individuals with pension income above minimum



Notes: The graph plots the evolution of life expectancy at 67 by pension quintile and sex, along with 95% confidence intervals. Estimates are constructed starting from a sample which includes only individuals with pension income above the minimum in each calendar year. Own elaboration based on INPS data.

Figure C.15: Lifetable entropy by pension income at 67 - individuals with pension income above minimum



Notes: The graph plots the evolution of lifetable entropy at 67 by pension quintile and sex, along with 95% confidence intervals. Estimates are constructed starting from a sample which includes only individuals with pension income above the minimum in each calendar year. Own elaboration based on INPS data.

Period life tables: estimation of Kannisto parameters

As specified in Section 3.3.3.3, we estimate the Kannisto parameters α and β through maximum likelihood by assuming that deaths follow a Poisson distribution with $D_x \sim Poisson(E_x \cdot \mu_x(\alpha, \beta))$, where D_x and E_x denote, respectively, deaths and person years lived at age x. We derive the Kannisto parameters by maximizing the log-likelihood function:

$$logL(\alpha, \beta) = \sum_{x=75}^{85} \Big\{ D_x log[\mu_x(\alpha, \beta)] - E_x \mu_x(\alpha, \beta) \Big\} + constant \tag{5}$$

where μ_x denotes mortality rate at age x.

D Appendix Chapter 5

D.1 Estimation of intercensal population in Italian provinces, 1953-1981

We collect and digitize data on: (i) census population by province, single-age class (0, 1, ..., 100+), and sex for years 1951, 1961, 1971 and 1981, (ii) yearly births count by sex and province, and yearly deaths count by province, sex and age class (0, 1, 5, 10,..., 95, 100+) for years 1953-1981. To reconcile deaths and population data, we first disaggregate yearly deaths counts into 1-year age groups fitting a univariate penalized composite link model, which assumes that death counts are Poisson distributed and that the underlying sequence on a fine grid to be estimated is smooth. The operation is performed through the 'ungroup' R package developed by Pascariu et al. (2018) for each province and year, for men and women separately. We then derive intercensal population estimates in each province considering that population in each intercensal year is the result of four main demographic processes: births, deaths, ageing and migration. For this purpose, we adapt the methodology developed by the Italian National Institute of Statistics (1985) to reconstruct regional-level population for years 1972-1981. We face three main data limitations. First, we do not have access to province level yearly births and deaths count for years 1951 and 1952, which are needed for intercensal population estimates. We address this issue by collecting and digitizing yearly births count by sex and region, and yearly deaths count by region, sex and age class (0, 1, 5, 10,..., 95, 100+) for years 1951 and 1952.[18] We then distribute births and deaths in 1951 and 1952 across provinces within each region depending on province-specific weights of births and deaths in 1953. In other words:

$$Deaths^{t}_{r,p,s,x} = Deaths^{t}_{r,s,x} \times weight^{1953}_{r,p,s,x} = Deaths^{t}_{r,s,x} \times \frac{Deaths^{1953}_{r,p,s,x}}{Deaths^{1953}_{r,s,x}}$$

[18]In Italy, regions represent the upper administrative tier, encompassing provinces (intermediate tier) and municipalities (lower tier).

$$Births^{t}_{r,p,s} = Births^{t}_{r,s} \times weight^{1953}_{r,p,s} = Births^{t}_{r,s} \times \frac{Births^{1953}_{r,p,s}}{Births^{1953}_{r,s}}$$

where subscripts t, r, p, s and x indicate year, region, province, sex and age respectively, and $t \in (1951, 1952)$. Second, we do not have access to data about migration inflows and outflows for Italian provinces in the years at stake. We tackle this issue by assuming that cumulative net migration flows between any two censuses are equivalent to the difference between observed and estimated population figures at the time of the second census, and that net migration flows distribute uniformly over any intercensal decade. Third, we do not have census population data by year of birth, but only by 1-year age classes, from age 0 to the open age-class 100+. With these caveats in mind, we derive intercensal population estimates in each province for ages 0 to 89, for men and women separately, taking the following steps 1 to 5. We use the intercensal decade 1952-1961 to illustrate the procedure.

Step 1. We estimate population on January 1, 1952 based on 1951 census population, adjusted for births, deaths and ageing occurring between 1951 census day (November 4) and January 1, 1952. We do not have data about daily births and deaths. Thus, we assume that births and deaths at age x occurring between 1951 census day and January 1952 are equivalent, respectively, to $\frac{d}{365} \times Births_{1951}$, in the case of births, and $\frac{d}{365} \times Deaths_{x,1951}$, where d represents the number of days between 1951 census day and January 1, 1952, while $Births_{1951}$ and $Deaths_{x,1951}$ are the total number of births and deaths at age x in 1951. As we do not have population census data by year of birth, we assume that individuals turn older uniformly over the year. In other words, we assume that $\frac{d}{365}$ of population aged x on 1951 census day will turn *x+1* by January 1, 1952, where *Popcens* is the population aged x on 1951 census day. For $0 < x \leq 99$, population on January 1, 1952 will thus be given by:

$$\overline{Pop}_{x,1952} = Popcens_{x,1951} + Popcens_{x-1,1951} * \frac{d}{365} - Popcens_{x,1951} * \frac{d}{365} - Deaths_{x,1951} * \frac{d}{365} \tag{6}$$

For $x = 0$, population on January 1, 1952 will be given by:

$$\overline{Pop}_{0,1952} = Popcens_{0,1951} + Births_{1951} * \frac{d}{365} - Pop_{0,1951} * \frac{d}{365} - Deaths_{0,1951} * \frac{d}{365} \tag{7}$$

Step 2. We estimate population on January 1, 1962 based on 1961 census population, adjusted for births, deaths and ageing occurring between 1961 census day (October 15) and January 1, 1962. We proceed as in step 1. For $0 < x \leq 99$, population on January 1, 1962 will be given by:

$$\overline{Pop}_{x,1962} = Popcens_{x,1961} + Popcens_{x-1,1961} * \frac{d}{365} - Popcens_{x,1961} * \frac{d}{365} - Deaths_{x,1961} * \frac{d}{365} \tag{8}$$

For x=0, population on January 1, 1962 will be given by:

$$\overline{Pop}_{0,1962} = Popcens_{0,1961} + Births_{1961} * \frac{d}{365} - Pop_{0,1961} * \frac{d}{365} - Deaths_{0,1961} * \frac{d}{365} \tag{9}$$

where d represents the number of days between census day in 1961 and January 1, 1962.

Step 3. We estimate intercensal population on January 1 of years 1953-1962. We follow an approach analogous to that of Steps 1 and 2. Specifically, we assume that every individual aged x on January 1 of year $1953 \leq t \leq 1962$ will turn one year older by December 31 of year t, except those who die in year t. Assuming uniform distribution of deaths for each cohort across the years, for $0 < x \leq 99$ population on January 1 of year t will be given by:

$$Pop_{x,t} = Pop_{x,t-1} - \frac{1}{2} Deaths_{x,t-1} \tag{10}$$

where $Pop_{x,t-1}$ and $Deaths_{x,t-1}$ are, respectively, the intercensal population and the number of deaths at age x in year *t-1*. For x=0, population on January 1 of year t will be given by:

$$Pop_{0,t} = Births_{t-1} - \frac{1}{2} Deaths_{0,t-1} \tag{11}$$

where $Births_{t-1}$ and $Deaths_{0,t-1}$ are, respectively, the number of births and the number of deaths at age 0 in year *t-1*.

Step 4. We calculate the intercensal residuals $\Delta_{x,t}$, for years $t \in (1952, 1962)$, accounting for cohorts born prior and after the 1951 census and assuming that residuals distribute uniformly over the intercensal decade. Cohorts born prior to 1951 will be aged 11 or older by December 31, 1961. Considering that the open age interval is set at 100+, we are able to fully derive intercensal residuals for ages $11 \leq x \leq 89$ applying the following formula:

$$\Delta_{x,t} = \frac{(t - 1952) + \frac{d}{365}}{10 + \frac{d}{365}} \left[Pop_{x,1962} - \overline{Pop}_{x,1962} \right] \tag{12}$$

where $\frac{d}{365}$ represents the year fraction following the 1951 census, $Pop_{x,1962}$ is the population aged x in 1962 computed in Step 3 and $\overline{Pop}_{x,1962}$ is the population aged x in 1962 computed in Step 2.

Cohorts born between 1951 and 1961 will be aged between 0 and 10 on December 31, 1961. Assuming that births distribute uniformly over the years 1952-1961, for $0 \leq x \leq 10$ we are able to derive intercensal residuals by applying the following formula:

$$\Delta_{x,t} = \frac{x + 0.5}{x - (t - 1952) + 10.5}\left[Pop_{x,1962} - \overline{Pop}_{x,1962}\right] \tag{13}$$

where 0.5 embodies the assumption that births are distributed uniformly over years 1952-1961, $Pop_{x,1962}$ is the population aged x in 1962 computed in Step 3 and $\overline{Pop}_{x,1962}$ is the population aged x in 1962 computed in Step 2.

Step 5. For ages $0 \leq x \leq 89$, we adjust intercensal population estimates for years $t \in (1952, 1961)$ derived in Step 3 with the intercensal residuals derived in Step 4:

$$P_{x,t}^{intercens} = \overline{P}_{x,t} + \Delta_{x,t} \tag{14}$$

For intercensal decades 1962-1971 and 1972-1981, we follow the same procedure detailed in Steps 1 to 5. The operation is repeated for each province, for men and women separately.

D.2 Estimation of age-specific mortality rates in Italian provinces, 1953-1981

For ages $0 \leq x \leq 89$, we compute province-specific mortality rates, $m_{x,t}^{p}$, in years 1953-1961, dividing the number of deaths $D_{x,t}^{p}$ by mid-year population $E_{x,t}^{p}$, using population estimates derived as in Section D.1.[19]. To derive mortality rates for $90 \leq x \leq 100$, we apply the Kannisto model (Thatcher et al. 1998). According to the Kannisto model, mortality at older ages can be approximated as follows:

$$m_x = \frac{\alpha e^{\beta x}}{1 + \alpha e^{\beta x}} \tag{15}$$

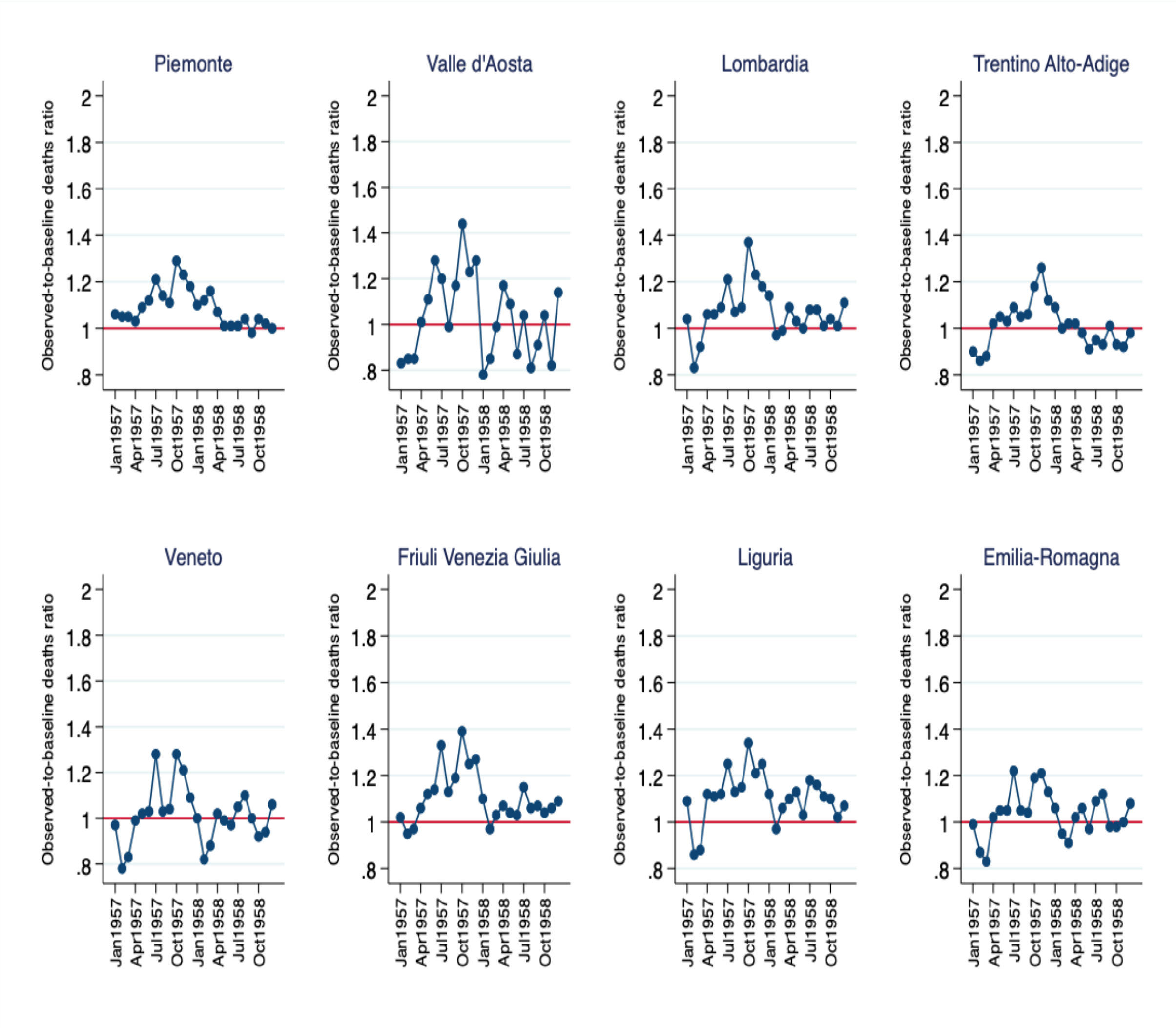
We fit the Kannisto model through ages 75-89 and estimate parameters α and β through maximum likelihood for each province, sex and calendar year combination, and use the estimated Kannisto parameters to extrapolate mortality rates from 90 to 100.

[19]For each year, $E^p_{x,t} = \frac{P^p_{x,t+1} - P^p_{x,t}}{2}$

Supplementary Figures

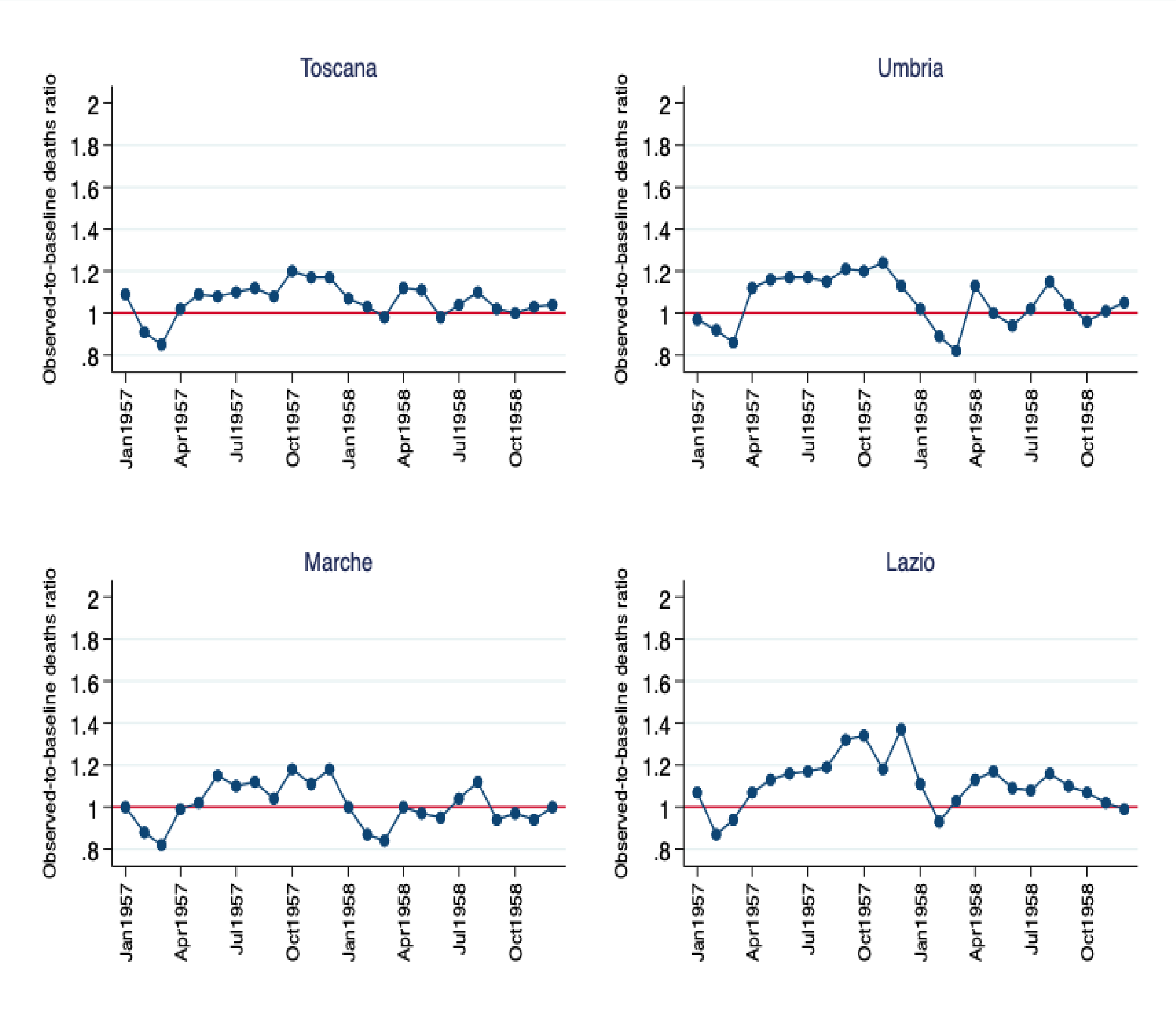
Figure D.1: Map of Italian provinces

Figure D.2: Ratio between observed and baseline all-cause monthly deaths in 1957-1958 in Italian regions, North



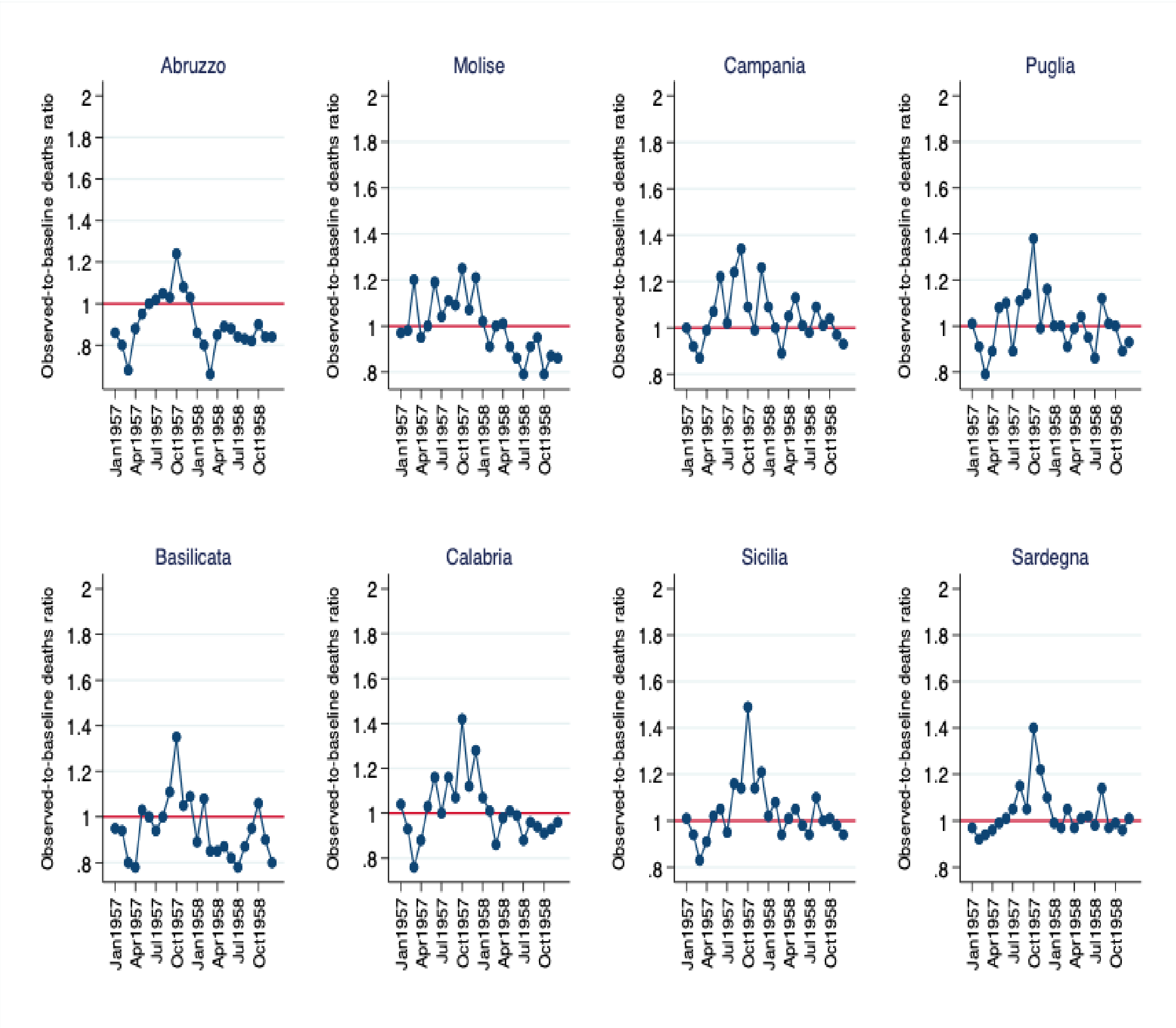
Notes: For each region-month of 1957-1958, the baseline is computed as the 1953-1956 average of total deaths. Own elaboration based on ISTAT data.

Figure D.3: Ratio between observed and baseline all-cause monthly deaths in 1957-1958 in Italian regions, Centre



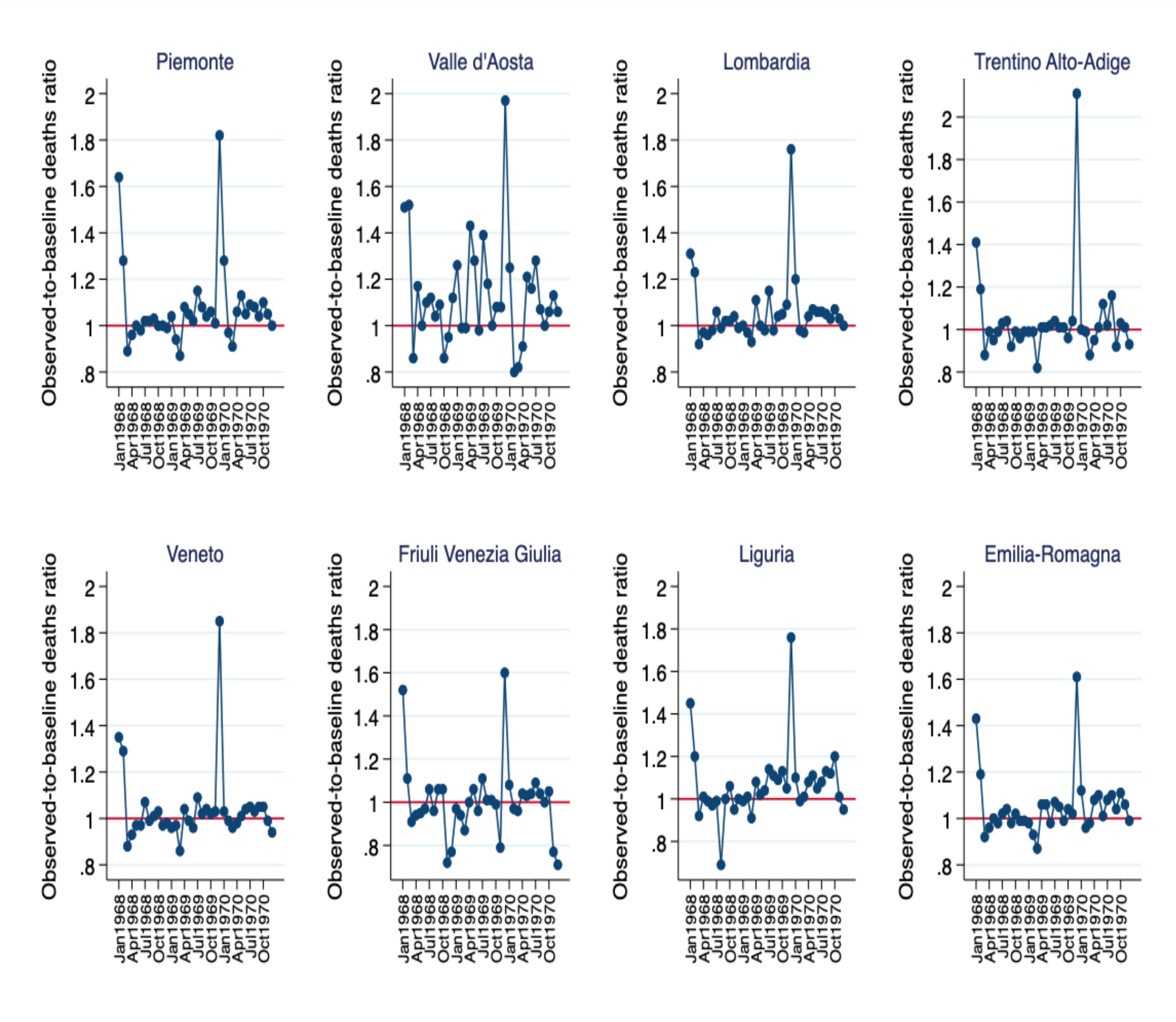
Notes: For each region-month of 1957-1958, the baseline is computed as the 1953-1956 average of total deaths. Own elaboration based on ISTAT data.

Figure D.4: Ratio between observed and baseline all-cause monthly deaths in 1957-1958 in Italian regions, South



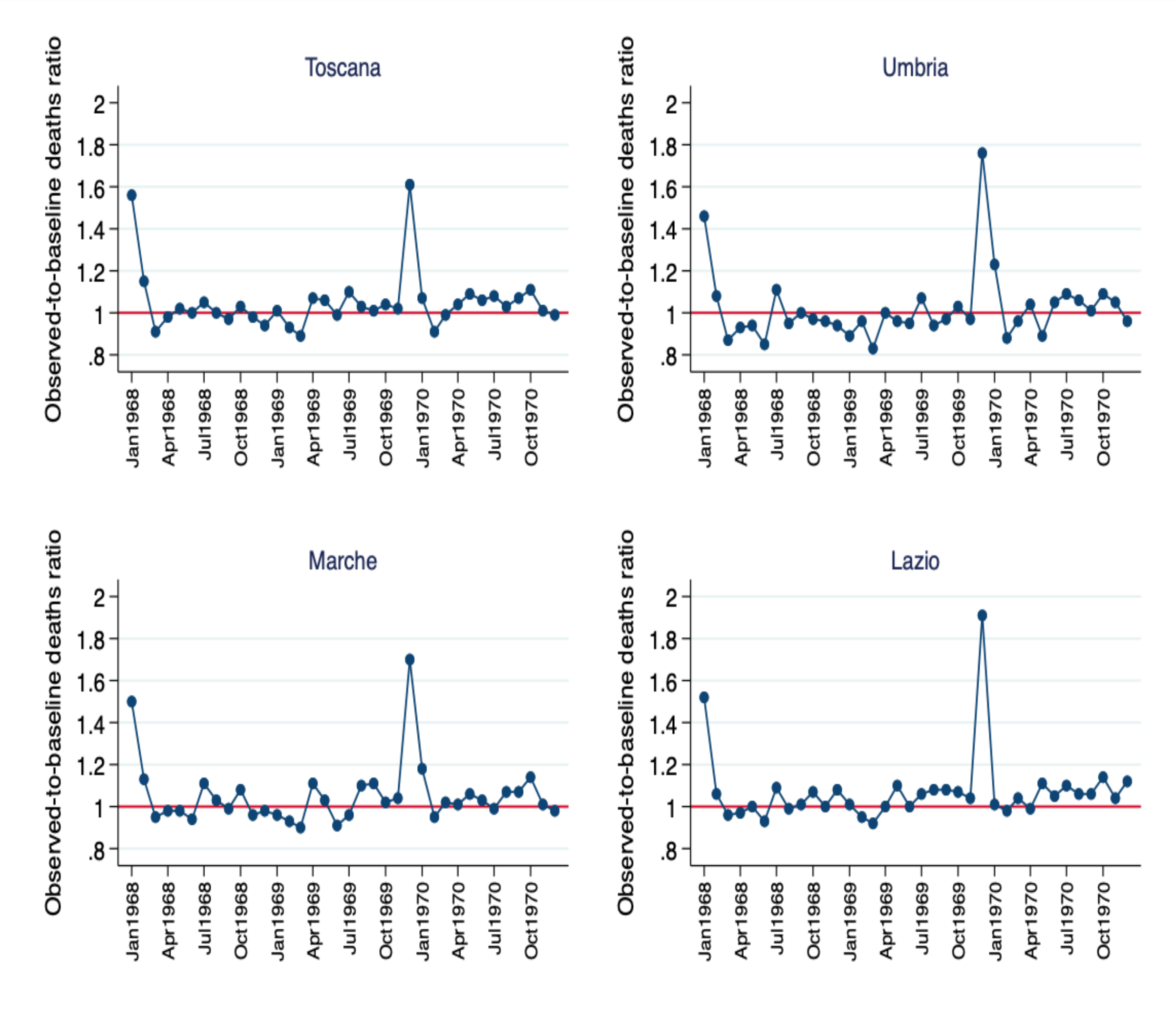
Notes: For each region-month of 1957-1958, the baseline is computed as the 1953-1956 average of total deaths. Own elaboration based on ISTAT data.

Figure D.5: Ratio between observed and baseline all-cause monthly deaths in 1968-1970 in Italian regions, North



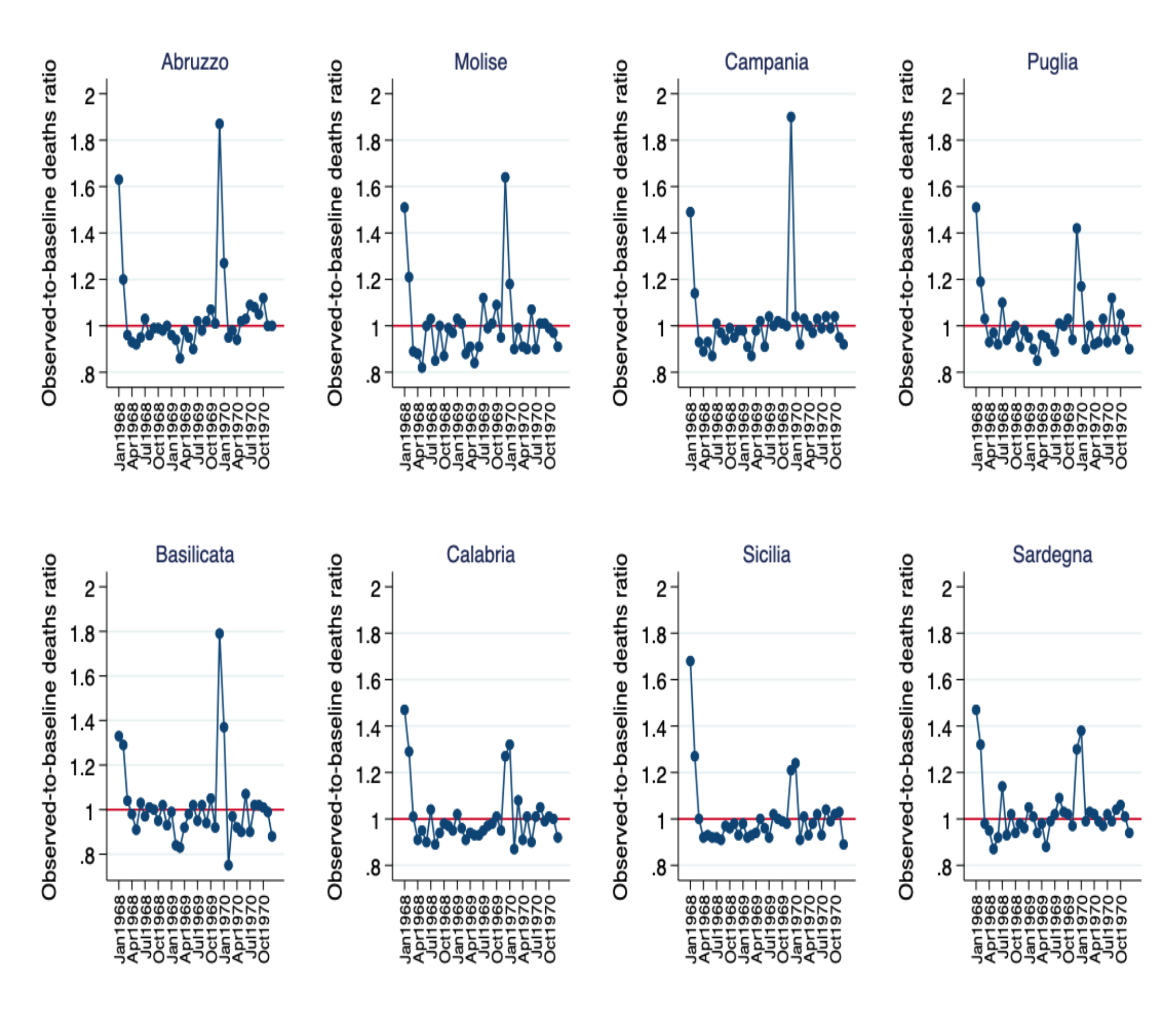
Notes: For each region-month of 1968-1970, the baseline is computed as the 1965-1967 average of total deaths. Own elaboration based on ISTAT data.

Figure D.6: Ratio between observed and baseline all-cause monthly deaths in 1968-1970 in Italian regions, Centre



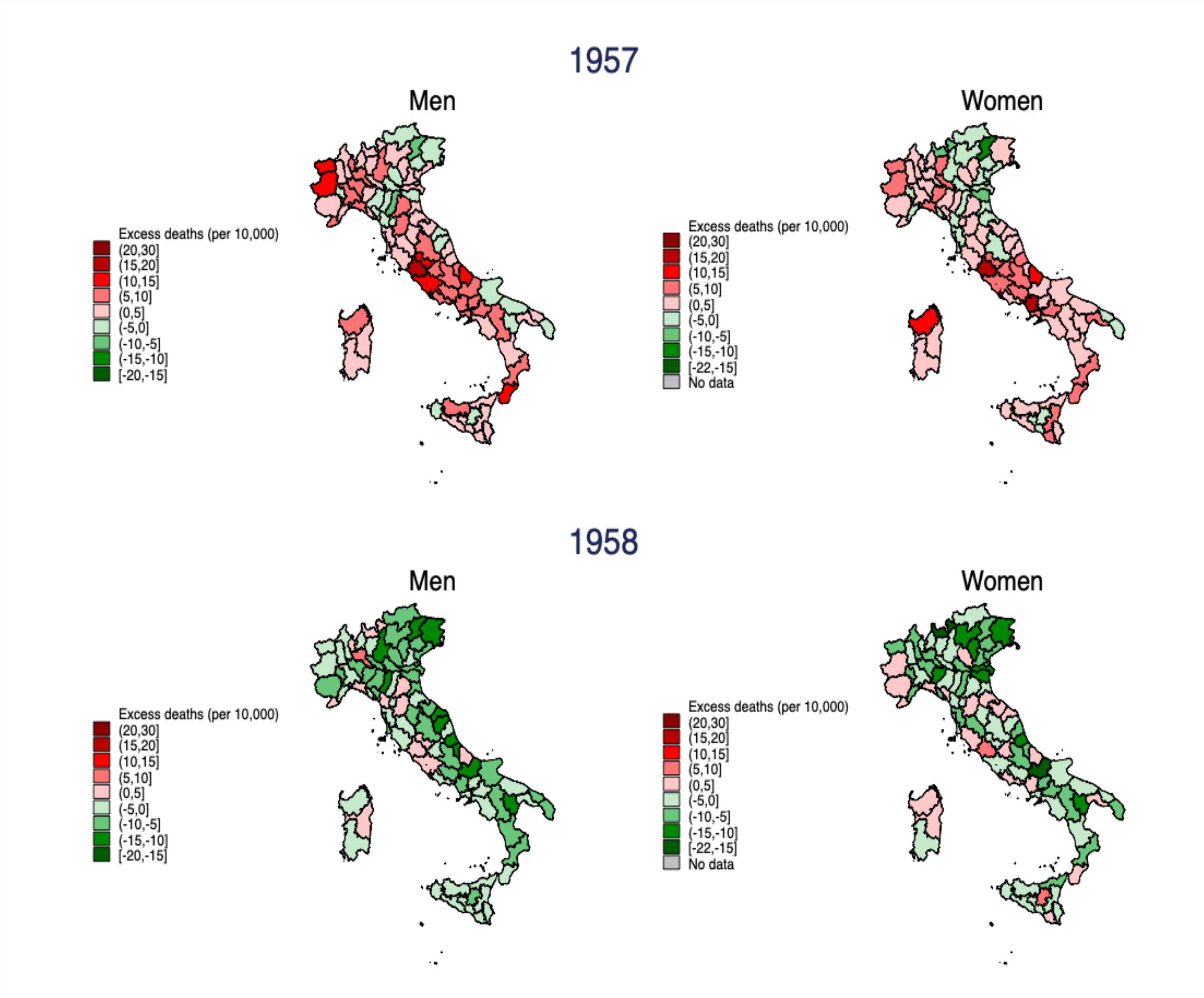
Notes: For each region-month of 1968-1970, the baseline is computed as the 1965-1967 average of total deaths. Own elaboration based on ISTAT data.

Figure D.7: Ratio between observed and baseline all-cause monthly deaths in 1968-1970 in Italian regions, South



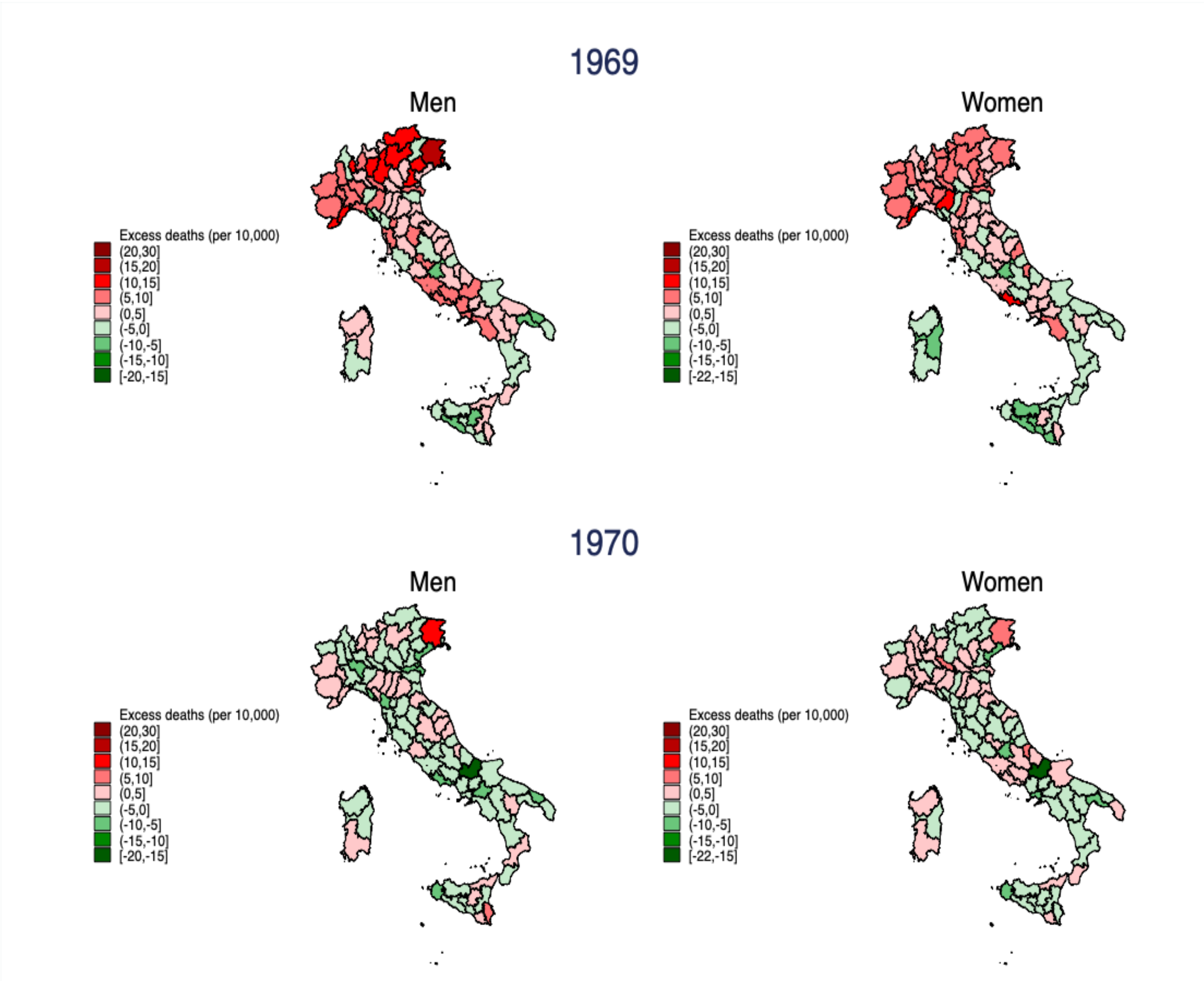
Notes: For each region-month of 1968-1970, the baseline is computed as the 1965-1967 average of total deaths. Own elaboration based on ISTAT data.

Figure D.8: Excess mortality per 10,000 in 1957-1958



Notes. Excess mortality is calculated as the difference between observed and baseline age-standardized mortality rates, per 10,000. In each year, standardization across provinces was performed using the Italian population as standard.

Figure D.9: Excess mortality per 10,000 in 1969-1970



Notes. Excess mortality is calculated as the difference between observed and baseline age-standardized mortality rates, per 10,000. In each year, standardization across provinces was performed using the Italian population as standard.