Hidden Redistribution in Lifetime Earnings: The Role of Differential Mortality*

Aurel Mélard† Simon Rabaté‡ Maxime Tô§

November 2025

Abstract

Public pension systems account for a significant proportion of public expenditure and involve substantial lifetime income redistribution, much of which is implicit and poorly documented. Using administrative data on private sector wage earners in France, we quantify and decompose this redistribution by comparing the return on contributions of different groups. Our findings show that the pension system is highly redistributive overall, with returns on contributions halving from the bottom to the top of the income distribution. This overall redistribution is largely driven by the progressivity of the pension formula, with differential mortality having a limited impact. However, this pattern masks substantial between and within gender redistribution. We observe significant redistribution from men to women, with women receiving an average return that is 33% higher than men's, mainly due to their higher life expectancy. Contrasted results are found within gender and between income groups: when accounting for differential mortality, progressivity for men is almost entirely reversed, while the system remains largely progressive for women.

*Corresponding author : Maxime Tô; maxime.to@ipp.eu; We thank Antoine Bozio, Patrick Aubert, Pierre Boyer, Carole Bonnet, Roméo Fontaine, Lamia Kandil, Jonathan Goupille-Lebret, Marion Leturcq, Jean-Baptiste Oliveau, Roland Rathelot, Johannes Spinnewijn, Julie Tréguier, Mickaël Zemmour and Gabriel Zucman, as well as participants to the PENSINEQ meetings for their valuable comments and suggestions. Remaining errors are our own.

†CREST, École Polytechnique, Institut polytechnique de Paris

‡Ined and Centraal planbureau

§Paris School of Economics, CREST and Institute for Fiscal Studies

1 Introduction

Over the last decade, research on income inequality has made enormous progress, examining the magnitude, evolution, and determinants of inequality between income groups (Piketty et al., 2018). There is also a better understanding of how tax and transfer systems attenuate these differences (Bozio, Garbinti, et al., 2024b, Bruil et al., 2022). Simultaneously, research on lifetime earnings inequality has advanced significantly (Kopczuk et al., 2010, Aaberge & Mogstad, 2015, Guvenen et al., 2022, 2021), revealing the importance of considering long-term economic trajectories. However, most studies focus on instantaneous redistribution, either at a single point in time or through repeated cross-sections. This approach hides two critical dimensions of lifetime redistribution. First, it fails to capture the intrinsically dynamic nature of some transfers, like education and social security. Second, it neglects differential mortality patterns: wealthier individuals and women tend to have longer life expectancy (Chetty et al., 2016), which has profound implications for lifetime income inequality and redistribution both within and between genders.

Public pension systems play a pivotal role in this context. Lifetime income redistribution is one justification for the existence of social security (Diamond, 2004). This typically takes the form of higher replacement rates received by individuals with low lifetime earnings. This explicit redistribution can be counterbalanced or amplified by implicit mechanisms driven by differential mortality. On the one hand, high-income individuals receive their relatively high pension benefits for longer periods, potentially exacerbating lifetime income inequality. On the other hand, women's pensions, while typically lower due to factors such as career interruptions for childbearing and caregiving, higher rates of part-time work, and occupational segregation (Goldin, 2014, Blau & Kahn, 2017, Kleven et al., 2019), are received over longer periods. This extended benefit period may reduce the gender gap on a lifetime basis. Given that pensions constitute the largest public expenditure in many countries, the effect of differential mortality on the progressivity of pension systems and overall tax and transfer systems is likely substantial. Despite the importance of this issue, there is a notable lack of evidence on how differential mortality shapes lifetime inequality and redistribution, both between genders and within genders across income groups. This paper aims to fill this gap in the literature by quantifying these effects and providing a comprehensive understanding of how differential mortality affects the substantial lifetime redistribution generated by public pension systems.

We analyse longevity differences by gender and income group, and their con-

sequences for lifetime earnings, using administrative records covering the entire population of private-sector wage earners in France. Our analysis yields three key sets of findings.

First, we analyse the relationship between income and life expectancy. Using permanent labour income, which is calculated as the average daily wage earned between ages 40 and 54, we compute life expectancy at 55 by income percentile for men and women separately. For both genders, we observe a clear income gradient in life expectancy that is almost linear in income percentile and concave in income level. This gradient is more pronounced for men (a seven-years difference between first and last percentiles against an average population life expectancy at 55 of 27 years) than for women (a two-years difference against an average of 33 years).

Second, we compute the overall income redistribution generated by the pension system using the return on contributions (RC) indicator. The RC compares benefits received from the pension system to contributions paid over the life course and has a direct interpretation in terms of redistribution: individuals or groups receiving above average RC are net beneficiaries from the system, while those receiving below-average RC are net contributors. We calculate the RC by ventile of lifetime earnings and find that the pension system is highly redistributive, with returns on contributions decreasing significantly with income. The RC for the lower income group is almost twice that for the highest group (180% vs. 100%). We further decompose the RC between two components: redistribution generated by the pension formula and redistribution attributed to differential life expectancy. This decomposition reveals that the overall redistribution is largely driven by the progressivity of the pension formula, with differential mortality having limited impact at the aggregate level.

Thirdly, we analyse the RC by gender and demonstrate that the patterns of aggregate redistribution are primarily explained by the gender composition of income groups. Specifically, we find that women are overrepresented at the bottom of the income distribution while men dominate at the top. Since women exhibit consistently higher returns on contributions across all income levels – driven by both longer life expectancy and more favourable treatment under solidarity components of the pension formula – the aggregate progressive pattern observed when pooling both genders largely reflects this compositional effect rather than genuine within-gender redistribution. Overall redistribution thus masks significant and contrasting redistribution between and within genders.

Between genders, we observe significant redistribution from men to women, with the average RC for women being 35 percentage points (33%) higher than for men.

Two thirds of this difference is explained by women's higher longevity, with a remainder due to the pension formula. Since the RC indicator does not directly reflect standard of living, we analyse the gender gap in lifetime pensions. We find that, absent differences in longevity, the gender gap would be twice larger than the observed one (27% vs. 13%).

Within gender and between income groups, we find that differential mortality has a significant impact on men. Among men and ignoring differentiated mortality, the RC decreases with income due to the solidarity components of the pension formula. Remarkably, accounting for differential mortality almost entirely reverses the RC progressivity for men, as the RC increases with income up to the last quartile of the income distribution. For women, differential mortality does not significantly affect redistribution, due to the combination of smaller estimated mortality gradient and substantial solidarity effects that strongly shape the RC by income group.

Thus, our findings reveal that differential life expectancy is the primary mechanism driving redistribution from men to women, while simultaneously undermining progressive redistribution among men by favouring higher-income individuals.

This paper contributes to three strands of the literature.

First, we connect to the literature on the determinants of life expectancy, particularly the relationship between income and mortality. Building on the seminal work of Chetty et al. (2016), which uncovers a strong correlation between household income and life expectancy from large administrative tax records in the United States, we propose alternative methodological choices.[1] Like Haan et al. (2020) and Milligan & Schirle (2021), we use a measure of permanent income to rank individuals, avoiding potential biases that may arise from income mobility (Kreiner et al., 2018). We also focus on individual labour earnings rather than household income. Labour earnings constitute the tax base for social security contributions and are therefore arguably the relevant income concept for analysing redistribution in pension systems. However, labour earnings may be a less accurate predictor of life expectancy, since they are not perfectly correlated with standard of living, especially for women. We provide a discussion regarding the choice of income concept and its practical consequences on the estimation of life expectancy gradient.

Second, we contribute to the literature on the distributive effects of pension systems, particularly their interaction with differential mortality. While static analyses of pension progressivity are common (Whitehouse, 2006, Aubert, 2012), evidence on lifetime redistribution is more scarce as it requires comprehensive data on income

[1] Chetty et al. (2016), has been replicated in different countries: Kinge et al. (2019) for Norway, Kreiner et al. (2018) for Denmark, Blanpain (2019) for France, and Hagen et al. (2025) for Sweden.

and mortality trajectories. Most studies overcome this constraint by relying on dynamic microsimulation models (Liebman, 2002, Mazzaferro et al., 2012, Dubois & Marino, 2015), in which individual trajectories are simulated based on estimated processes (Li & O'Donoghue, 2013). These studies typically find that public pension systems are redistributive – providing higher return on contributions for low socioeconomic status individuals – and identify only moderate effects of differential mortality on overall and within gender redistribution. However, these studies are limited by their reliance on modelling assumptions as well as their use of survey data and education-based mortality estimates. They cannot leverage the full range of relationships between income and life expectancy that were demonstrated by Chetty et al. (2016) and subsequent studies. Haan et al. (2020), overcomes these limitations by using administrative data containing complete individual trajectories to estimate life expectancy and returns on contribution by groups of lifetime income. Consistent with our finding, they demonstrate that redistribution in the public pension system for men in Germany is reversed once differential life expectancy is taken into account. An important limitation of this paper is its focus on men only, with no analysis of overall or between gender redistribution within the system. Focusing on Sweden, Shi & Kolk (2022) consider both men and women, but they only consider them as separate populations, and focus on lifetime pensions inequality without accounting for contributions, providing an incomplete assessment of redistribution. To our knowledge, our paper is the first to provide a comprehensive evaluation of lifetime redistribution in the public pension system using microdata. By combining the general scope of analyses based on a microsimulation models with the accuracy of exhaustive administrative data, we are able to quantify and decompose the various types of redistribution generated by the pension system.

Third, we contribute to research on gender gaps in pensions. This literature typically focuses on cross-sectional comparisons of pension levels (Betti et al., 2015, Samek Lodovici et al., 2016), examining factors such as gender differences in work hours (Bertrand, 2020, Goldin & Lleras-Muney, 2019) and hourly wages (Blau & Kahn, 2017), which translate into lower pension rights. There are also differences between men and women in terms of savings behaviour (Neelakantan & Chang, 2010) and retirement decisions (Tréguier, 2021), which can be related to differences in risk aversion or social preferences (Bertrand, 2011). Our study adds to this literature by quantifying lifetime redistribution between genders in the pension system and assessing how women's higher longevity contributes to narrowing the pension gap relative to career-related factors.

The rest of this paper proceeds as follows. Section 2 describes the institutional

context and the data sources used in the analyses. The theoretical concepts and empirical methods we employ are detailed in Section 3. We then present the results on the relationship between income and mortality and on redistribution generated by the pension system in Section 4. Section 5 concludes.

2 Background

2.1 Institutional context

The public pension system in France is large, with benefits amounting to roughly 14% of GDP. In this paper, we focus on wage earners in the private sector. For this group, pensions are divided into two components that are provided by two different pension schemes: a first pillar (*Régime général*) and a second pillar (*Agirc-Arrco*). A high share of the French working population is affiliated with these regimes at least once in their careers: pensioners of the *Agirc-Arrco* scheme – on which we focus in this study – represent 72% of current and 84% of new pensioners in France in 2021 (DREES, 2023).

The first and second pillars share important common features: they are both pay-as-you-go pension schemes funded by compulsory social security contributions. Their main differences concern their contribution bases their pension formulas – annuities for the first pillar and a point system for the second – and their governance – the state for the first pillar and labour organisations for the second.[2]

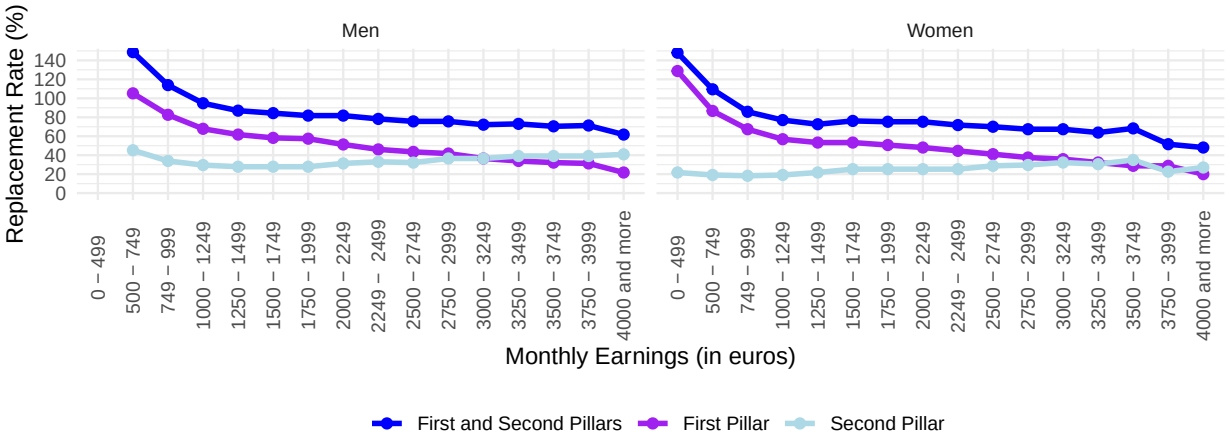
Both pillars are contributory, in the sense that pension benefits increase with pension contributions. However, the pension formula and the way solidarity components (accounting for births, disability, unemployment, etc.) are incorporated into the system may lead to important differences in returns on contributions. Figure 1, taken from Aubert et al. (2012), shows that the replacement rate (the ratio between pension and end of career labour earnings) is clearly decreasing with monthly earnings, suggesting that returns may be higher for low income workers.

This decreasing shape is mainly driven by the first pillar, which incorporates more solidarity components in its formula. In contrast, second-pillar replacement rates do not vary much according to labour earnings, reflecting the stronger contributory feature of the second pillar. Aubert & Bachelet (2012) show that the redistributive aspect of the first pillar is mainly driven by the solidarity components associated with parenthood or those that compensate for career interruptions. Women benefit particularly from this redistribution through solidarity components that are

[2]Online appendix A presents in detail the main features of the pension formulas for the two schemes.

specifically targeted to compensate for career interruptions linked to maternity and the education of children.[3] The basic pension formula is actually anti-redistributive: when neutralising all solidarity components, the D9/D1 ratio of pension is larger than the D9/D1 ratio of cumulated earnings.

Figure 1: Median replacement rate in the private sector, by scheme and income group



NOTE: This figure presents median replacement rates defined as the ratio between the pension at claiming age over labour earnings before retirement, as function of monthly labour earnings before pension claiming.
SOURCE: Aubert et al. (2012)

2.2 Data sources

The primary data sources for this study come from the administrative records of the *Agirc* and *Arrco* pension schemes. These provide detailed information on current rights and past contributions of individuals who were affiliated to the scheme – both retirees and workers – at some point between 2009 and 2022. Demographic data on these affiliates include individuals' dates of birth and death, gender, and country of origin.

This data allows us to recover individuals' earnings trajectories in the private sector from 1962 – creation of *Arrco*, the sub-scheme for non-executives – to 2022. Specifically, we have information on earnings and points accumulated each year by affiliates for different types of points (directly through contributions or non-contributory points for unemployment, illness, or disability). In addition to career information, *Agirc-Arrco* records provide exhaustive information on second-pillar pensions, with extensive details on the different components of these rights – direct

[3]These components are detailed in appendix A.4.

rights from contributions but also non-contributory rights from each specific solidarity schemes. We also have information on the benefits received form all other pension schemes for retirees after 2014.

The main advantage of this administrative data source is that it enables us to observe the complete career and retirement trajectory of affiliates. This makes it possible to analyse redistribution from a lifetime perspective.[4]

Other data sources used in our empirical analysis include series of mortality by sex, age and year for the overall French population, from the National Statistics Institute – *Insee*[5] – and various other sources to recover historical parameters of interest for the pension system. These are detailed in online appendix B.1.

3 Concepts and methods

The empirical analyses in this paper consist of two main steps: the estimation of an income gradient in life expectancy and the assessment of overall redistribution in the pension system, in light of these life expectancy differences between and within genders. In this section, we present the methodology developed in each analysis. First, we present the income measure that we use in both analyses to rank individuals in the income distribution.

3.1 Income definition

Income concept The literature studying the relationship between income and mortality does not use a uniform and consistent measure of income to rank individuals. Chetty et al. (2016) and most followers rank individuals based on their instantaneous household income. We depart from this definition in two ways.

First, we define individuals' income rank based on their position in the income distribution over the life-cycle. Second, we use individual labour earnings instead of household income or standard of livings – as done in Chetty et al. (2016) and Blanpain (2019).

This choice is motivated by two reasons: on the one hand, individual labour earnings is the most relevant income concept to study lifetime redistribution in the pension system, as we relate total amounts of contributions paid on labour income to total amounts of pensions received. On the other hand, using permanent income

[4]Note that data quality improved over time, especially regarding career data: in online Appendix B.2, we show that data on career trajectories is less complete for less recent years, individuals who retired before 1999, and more generally for older cohorts.

[5]See https://www.insee.fr/fr/statistiques/7678712?sommaire=7678731

rather than lagged income as a determinant of mortality potentially limits the issues raised by income mobility during the life cycle (see Kreiner et al., 2018).

Ideally, we would like to rank all individuals using total labour earnings over the life-cycle. In practice, we depart from this ideal setting by focusing on a measure of "permanent income", defined as average labour income between ages 40 and 54. This choice is driven by sample restrictions stemming from conceptual and data constraints that we explain below.

Sample selection The population we focus on results from a trade-off between our measure of permanent income and our measure of life expectancy: on the one hand, we want our measure of permanent income to approximate lifetime earnings. On the other hand, we want to have the largest sample possible, especially at older ages, to compute life expectancy at 55.

The upper bound of the age window for our income measure is defined to avoid reverse causality between income and mortality. Since we measure mortality from age 55 onward, we choose 54 as the upper bound of the permanent income measure.

The lower bound results from a trade-off between observing deaths at older ages and the fact that for older cohorts, we only observe careers from 1962 onward. We would need to focus on individuals born after 1945 to observe their full working career. These cohorts are still relatively young – at most 74 – in 2019, the last year selected for our mortality analysis. This imposes strong constraints for the estimation of life expectancy. We end up considering cohorts born between 1922 and 1964. For these, we can compute a permanent income between 40 and 54 and observe mortality from 55 to 90. Among these cohorts, our initial sample comprises individuals who were alive in 2009 and at age 55 or older.

On this initial sample, we want to focus on affiliates for whom our measure of permanent income is deemed representative of the total career. This leads us to make further sample restrictions. First, we focus on France-born individuals: if a large share of the career is spent abroad, our measure or earnings might not reflect actual lifetime labour income. Mortality is also poorly measured for this population, when these affiliates do not live in France anymore at time of death. Second, we restrict the sample to individuals for whom we observe at least 3 years of contribution between ages 40 and 54. Below this threshold of time spent in the *Agirc-Arrco* scheme, we consider that private sector income is likely to be poorly correlated with total lifetime income. For the same reason, we exclude the bottom 7% of the sample. Among individuals below this threshold for whom we observe all

pensions[6], we indeed observe that a relatively large share of career is often spent in other regimes (see Appendix Figure B.3), making our measure of income a poor predictor of mortality.

These restrictions have a significant impact on sample size.[7] Starting from 14.3 million men and 12.3 million women, we end up with 5.0 million men and 3.9 million women in our final sample. However, it should be noted that the important restrictions tend to exclude individuals who are not working much in the French private sector: they may be workers from other sectors (such as the public sector) who occasionally contributed to the private sector, or people out of the labour force between ages 40 and 54. In Appendix C.1, we find that our measure of life expectancy gradient is not much affected by alternative sample definitions.

Income rank computation We define permanent income as average observed earnings between ages 40 and 54. For a given year, earnings are either directly observed or derived from contributed points and individual contribution rates.[8] Earnings are then normalised by an index of mean wage per capita and summed over ages 40 to 54. Total earnings are divided by the observed number of days worked during the same period to yield our measure of permanent income.

We then rank individuals into percentiles of permanent income. Since mortality of some cohorts are only observed at ages older than 55, simply defining quantiles of permanent income based on our observed sample would be exposed to the risk of survival bias (Attanasio & Hoynes, 2000): among survivors of older cohorts, individuals with high levels of income are likely to be overrepresented. We deal with this issue by defining percentiles of permanent income based on the distribution of our younger cohort – 1964 – for which all survivors at age 55 (the younger age of the mortality analysis) are observed. Percentiles are defined with men and women pooled. The resulting income thresholds are used to rank all individuals from older cohorts into these percentiles.

This definition of permanent income ranks relies on the following main assumptions. First, we assume that mean observed earnings per day worked between ages 40 and 54 is a good predictor of total earnings during this period. In that respect, the way we treat periods of work interruption or work outside of the private sector matters. By dividing observed earnings by the number of days worked, we implic-

[6]Those constitute a subsample of the study, who did not die before 2014 or before the age of claiming.

[7]Appendix Table A.1 summarises the impact of each restriction on sample size.

[8]Point data is more complete than earnings data in the career records. Individual contribution rates are inferred from periods where both earnings and points are observed. The precise methodology is described in Appendix B.1

itly assume that individuals' earnings during unobserved periods were similar to the ones observed during work periods in the private sector.

Second, we assume that earnings between ages 40 and 54 are a good predictor of lifetime earnings. Some of the sample restrictions mentioned above aim at making this assumption reasonable: the selected sample comprises individuals for whom working periods in the private sector during ages 40 and 54 represent a sizeable share of their whole career.

Third, we assume that across cohorts, individuals with the same level of permanent income are similar – in particular regarding mortality patterns. Since we normalise earnings by an index of mean wage, we control for wage growth over time. Note that we do not make assumptions on the evolution of the income distribution across cohorts. Income groups correspond to percentiles of the 1964 cohort but are not bound to represent 1% of each cohort. We therefore implicitly assume that what predicts mortality, rather than the rank in one's cohort distribution, is the level of earnings, relative to the prevailing mean wage during one's career.

In Appendix C.1, we test alternative definitions of permanent income and permanent income ranks and observe the effect on the observed life expectancy gradient. These alternative measures aim at testing these three assumptions. We find that our measure of life expectancy gradient is robust to these sensitivity analyses.

3.2 Measurement of life expectancy gradients

Mortality is defined using cross-sectional data on observed deaths and survival of cohorts 1922-1964, averaged over the 2009-2019 period. Note that in our data, we observe mortality up to 2022. However, because of excess mortality linked to the Covid epidemics, we exclude 2020 and subsequent years. Observed mortality rates at age a for a certain sub-sample x – a subset of the overall sample based on income group or gender – are defined in cross-section, averaged over the 2009-2019 period:

$$\hat{Q}_{xa} = \frac{\sum_{t\in\{2009,2019\}} D_{xat}}{\sum_{t\in\{2009,2019\}} N_{xat}}$$

Where N and D are respectively the number of survivors as of January 1 st of the considered year and the number of deaths during the year. Age is that which would be reached during the year.

As in most of the aforementioned literature on mortality inequality, we use a parametric model of mortality to estimate mortality rates Q_{xa} from observed $\hat{Q}_{xa}$. We assume that the logarithm of mortality coefficients evolves quadratically with age from 55 to 90. More precisely, to deal with zero values in observed mortal-

ity rates, we estimate the following Poisson Pseudo-Maximum Likelihood (PPML) model:

$$E(Q_{xa}) = \exp(\alpha_x + \beta_{1x}a + \beta_{2x}a^2) + \varepsilon_{xa}, \quad E(\varepsilon_{xa}|a) = 0$$

Smoothed mortality rates at age a, $\tilde{Q}_a$, are obtained as the predicted values of the model. For ages above 90 and up to 120, mortality rates are those of the overall French population of the same gender. Survival at age 54 is equal to 1 by definition and from age 55 to 120:

$$\forall\ a \geq 55\ :\ S_{xa} = \prod_{i=55}^{a}(1 - \tilde{Q}_{xi})$$

Life expectancy at 55 is equal to the sum of survival probabilities at all subsequent ages ($S_{x54} = 1$ and $S_{x121} \approx 0$):

$$\begin{aligned} E_{x55} &= \frac{\sum_{k=55}^{120}(k-54)(S_{xk} - S_{xk+1})}{S_{x54}} \\ &= \sum_{a=55}^{120} S_{xa} \end{aligned}$$

The way we model mortality relies on two main assumptions. First, we assume that the log of expected mortality coefficients can be adequately approximated by a quadratic function of age in the range that we study. This specification is close to established practices in estimating mortality gradients across socioeconomic groups, as demonstrated by studies examining life expectancy inequalities (Chetty et al., 2016). In Appendix C.2, we test main alternative specifications found in the literature. We find that life expectancy gradients between and within genders are mostly left unchanged.

Second, while instantaneous mortality is usually computed for a given year, in our analysis we aggregate 11 years, from 2009 to 2019, for the computation of a given mortality rate by age and subpopulation. We do so primarily for sample size reasons. The resulting observed mortality rates are thus a weighted average of annual mortality rates over these 11 years. One might be concerned that we place more weight on recent periods – more recent cohorts with larger sample sizes. One might also be interested in understanding the evolution of mortality patterns over time. Research has documented significant changes in life expectancy gradients in recent decades, with studies showing increasing inequalities within gender

groups and evolving patterns between genders (Chetty et al., 2016, Dahl et al., 2021, Hagen et al., 2025). In Appendix C.2, we estimate the same mortality model by year and show the evolution of the life expectancy gradient between 2009 and 2019: within gender, life expectancy inequalities increased, while between-gender inequalities tend to decrease.

3.3 Lifetime redistribution in the pension system

Return on contributions The main measure we use to analyse redistribution in the pension system is the return on contributions – henceforth referred to as RC. It is defined as the ratio of the lifetime flow of discounted pensions during retirement to the lifetime flow of discounted contributions during one's career. For a given individual i, we define it as:

$$RC_i = \frac{\sum_{a=R_i}^{120} \beta^a S_{ai} P_i}{\sum_{a=8}^{68} \beta^a C_{ai}} \tag{1}$$

where a is age, β is the discount factor[9] and P_i, R_i, C_{ai}, and S_{ai} respectively are the annual pension level, claiming age, annual social security contributions and survival probability function. Note that all amounts are deflated using a consumer price index. This is consistent with the current rule in France, whereby pension benefits are indexed to prices, meaning they remain constant in real terms after claiming a pension.

The higher the RC, the bigger the flow of received pensions, compared to paid contributions. At the aggregate level, the average RC of a given cohort, in a pay-as-you-go system will be largely determined by the growth rate of the wage bill, which is the contribution basis of the pension system (Samuelson, 1958). However, for a given cohort, we can observe large variations in RC between individuals. These variations are driven by three factors: mortality, replacement rate – ratio of pension level over lifetime contributions – and claiming age. Variations in claiming age and mortality influence RC by playing on pension duration. Replacement rate can also vary between individuals based on the non-contributory elements of the pension formula[10].

[9]Throughout this analysis, unless otherwise mentioned, we use $\beta = 0.98$ as our discount factor. Other discount factors are tested in Appendix C.3. We find that higher discount factors accentuate the effect of mortality differentials by putting more weight on old ages.

[10]This includes additional rights for period of unemployment or low employment, or advantages related to children. See online appendix A for a description of the existing mechanisms in France.

We interpret variations in RC relative to the population mean as an indicator of redistribution in the pension system. The RC indicator allows us to identify the main explicit or implicit "winners" and "losers" of redistribution, in terms of lifetime monetary transfers, compared to a benchmark of equalised returns on contributions for all.

Comparing observed returns to this benchmark implicitly relies on an equity norm that has limitations, as do all inequality indicators. It is indeed uncertain whether pension systems aim at this equalisation. On the one hand, most pension systems involve a substantial contributory element. On the other hand, objectives of solidarity and insurance against longevity risk *de facto* lead to unequal returns. What we try to do, using RC, is to quantify and distinguish the redistribution operated by these mingled objectives.

Social Security Wealth We complement the RC indicator by computing Social Security Wealth – *SSW*. SSW corresponds to the flow of lifetime pensions received, *i.e.* the numerator of RC:

$$SSW_i = \sum_{a=R_i}^{120} \beta^a S_{ai} P_i \tag{2}$$

We use this additional indicator for two reasons. First, it allows us to compute the monetary value of identified transfers. Second, it circumvents a limitation of RC: an individual may benefit substantially from redistribution, while having a relatively low standard of living at retirement – a relatively high RC and relatively low SSW. This measure is therefore particularly useful when analysing redistribution in monetary terms between men – relatively higher pensions and lower returns – and women. We consider RC and SSW as two complementary measures when quantifying and qualifying redistribution in the pension system.

Decomposition To analyse the sources of redistribution, we decompose the observed differential in RC or SSW across sub-populations K (gender, income), by constructing counterfactual scenarios. In these scenarios, we replace sub-population-specific quantities with population averages (Ω) for the three components of these indicators shown in equation (1): survival probability, replacement rate (pension level relative to lifetime contributions), and claiming age. The methodology for SSW follows the same approach and is detailed in Appendix B.4.

We construct two counterfactual scenarios. In the first scenario, we equalise only mortality rates to obtain counterfactual returns on contributions if popula-

tion K had the same life expectancy as the whole population (RC_K^{LE}). In the second scenario, we additionally equalise replacement rates and claiming ages. We combine both dimensions under a pension formula component. In spite of some non-linearities in the pension schedule, pension benefits are roughly proportional to the retirement age so that both are largely jointly determined. This second scenario therefore controls for elements contained in the pension formula, which mainly accounts for solidarity components but also for any redistributive elements that may exist in the link between contributions and the level pension benefits. We refer to the returns on contributions in this scenario as $RC_K^{LE,PF}$. By construction, RC is equalised across all sub-populations in this second scenario.

Importantly, counterfactual returns are constructed under the constraint that the overall generosity of the system remains constant across scenarios. Formally, counterfactual returns for any sub-population are built such that RC_Ω – the average RC observed in the whole population – remains unchanged across scenarios. In each scenario, pension levels of all sub-populations are adjusted by the same proportional factor to respect this constraint.

This constraint ensures that RC in the second counterfactual scenario equals the observed population average ($RC_K^{LE,PF} = RC_\Omega$). We can therefore decompose the RC gap into its two components:

$$RC_K - RC_\Omega = \underbrace{RC_K - RC_K^{LE}}_{Impact\ of\ LE} + \underbrace{RC_K^{LE} - RC_\Omega}_{Impact\ of\ pension formula}$$

We apply this methodology first to decompose the gender gap in RC (section 4.3), then to decompose the gap in RC within gender between income ventiles (section 4.4). Note that decomposition results depend on the order in which we remove observed differentials. We therefore test the sensitivity of our results to this ordering in Appendix C.3, and find that the results remain unchanged.

Sample selection Our analysis of the redistribution based on return on contribution requires having a complete view of both lifetime contributions and pension levels in the private sector. For that purpose, we focus on individuals of the 1954 cohort: they were young enough at the creation of the second pillar scheme for executives and non-executives in 1962, and old enough to have retired as of 2022, the last observation in our data. This allows us to recover both individual contributions and pension levels.[11] We restrict our analysis to affiliates who had retired before

[11]Pension level is the one observed at date closest to retirement. Contributions of the first pillar are recovered from earnings and contribution rates, while contributions of the second pillar are recovered from points and their purchasing price. Contributions for work periods where point

this date and for whom we can reconstruct the whole flow of career contributions, namely around 329,000 individuals, of which 45% are women.[12]

Note that in this analysis, we use survival rates estimated on the 1922-1964 cohorts, which we apply to the 1954 cohort. Ideally, one would prefer to use mortality of the studied cohort at all ages above 55. However, mortality of cohort 1954 is only observed up to 2021 – at age 67. We therefore prefer our measure, that covers all ages from 55 to 90, excluding years of the Covid epidemics. In Appendix C.2, we estimate mortality by year between 2009 and 2019 and show that overtime, gradients increase within- and decrease between genders. Since the 1954 cohort is a relatively young cohort of our sample, our results on redistribution for this generation are probably a lower bound within gender and an upper bound between genders.

4 Empirical results

4.1 Life expectancy gradients

Figure 2a and 2b respectively show life expectancy at 55 for men and women by permanent income percentiles.[13]

Male life expectancy gradient For men, we find a 7.2-year gradient in life expectancy across permanent labour income. Life expectancy generally increases with permanent income and is almost linear in income rank. Minimum life expectancy at 55 equals 22.8 years and is associated to the 2nd percentile. Life expectancy at the very top of the distribution is around 30.8 years, leading to a 8.0-years difference between these two extremes.

Female life expectancy gradient Among women, we find a much lower gradient of around 1.5 years. This result might be driven by lower mortality inequalities among women,[14] but it is also likely to be strongly driven by the chosen income concept: labour income is less likely to be a strong predictor of total household income for women. The linearity of the increase in income percentile is however still apparent.

---

or earnings data is missing are imputed. The methodology for this imputation is detailed in Appendix B.2.

[12] This additional sample restriction and its impact on sample size is described in Appendix Table A.1.

[13] Appendix Figure A.5 presents the same result by level of 2022 euros annualised income instead of percentiles. Life expectancy appears to be increasing and concave with income level, especially among men.

[14] Figure A.3 shows that age of death among women is more concentrated at some ages.

It is also worth noting that life expectancy at 55 for the first percentile of women is 31.7 years, and it exceeds the life expectancy observed for men with the higher permanent income.

Both genders pooled Finally, Figure 3 shows the same gradient for men and women combined. Average life expectancy is decreasing with income rank until the 39th percentile and increasing above. The maximum life expectancy gap is relatively low and reaches 3.2 years. This gradient results from three different effects: First, life expectancy of women is higher than the one of men and relatively stable across income ranks. Second, life expectancy among men strongly increases with income for most of the gradient. Third, the share of women within each rank is strongly declining with income, as shown in Figure 3b.

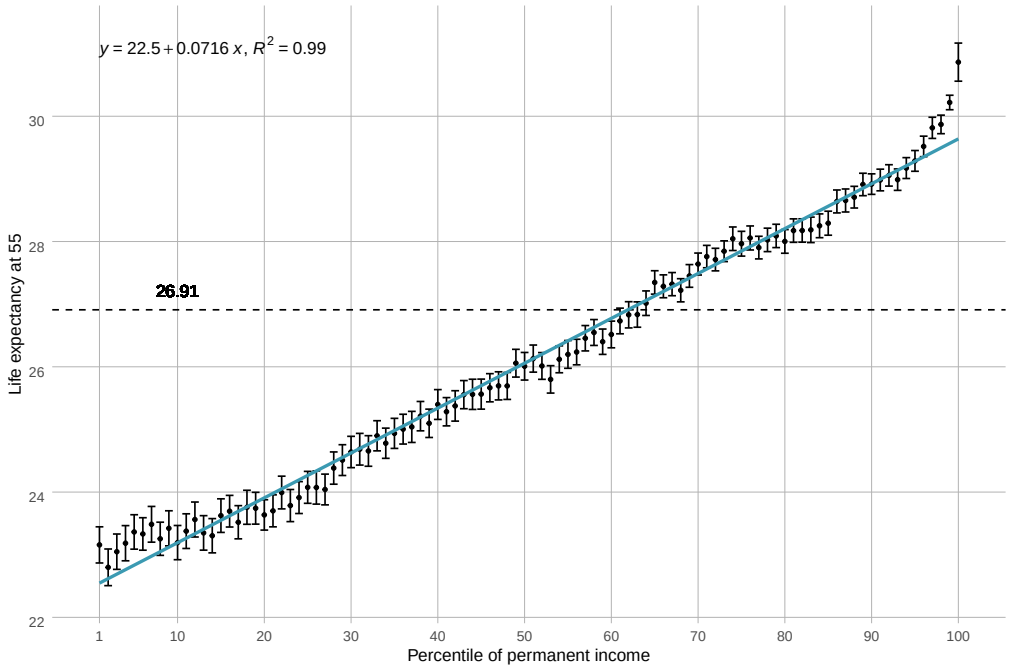
The observed gradient is the result of this composition effect between the gradient among men and that among women. This result has important redistribution implications, visible in section 4.2: when abstracting from between-gender variations in mortality, inequalities in life expectancy and subsequent implicit redistribution appear moderate.

Discussion on magnitude We use Blanpain (2019) as our primary point of reference, which shares methodological similarities with Chetty et al. (2016), and also focuses to the French population. This paper examines mortality between 2012 and 2016, using panel data following approximately 1% of all French residents. Consistent with Chetty et al. (2016), Blanpain uses an instantaneous measure of household earnings as the income metric. The income gradient of life expectancy at 55 between the top and bottom ventiles in Blanpain's study is more pronounced for men. She reports a 9.0 year difference compared to 7.2 in our study. The disparity is even more striking for women, with Blanpain finding a 6.2 year difference versus 1.5 in our study.

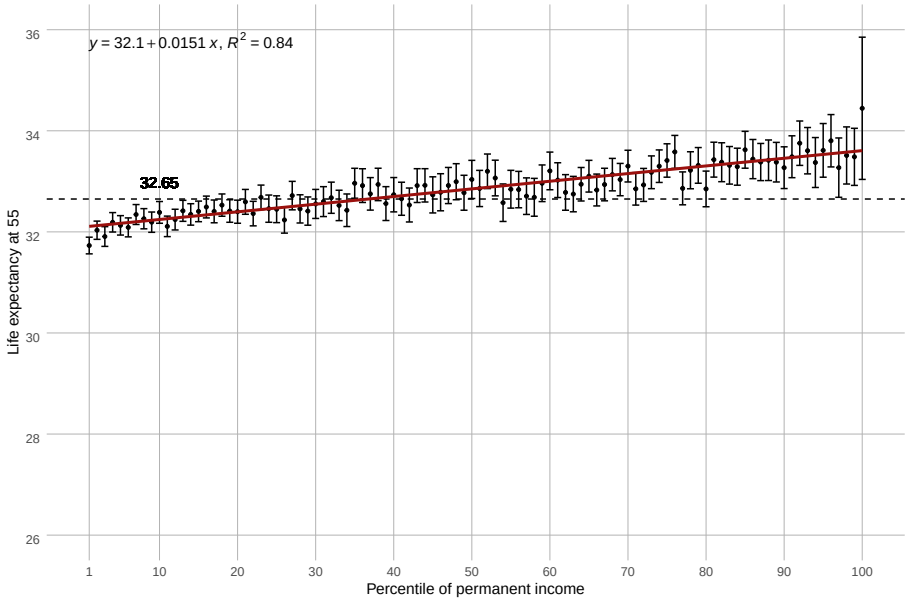
We attribute our comparatively lower gradients to three concurrent reasons. First, we exclusively focus on workers, whereas Blanpain (2019) and Chetty et al. (2016) consider the entire population. We therefore probably exclude some precarious, non-working segments of the population who may have lower life expectancy. Second, we use individual labour income as our metric, while Chetty et al. (2016) and Blanpain (2019) focus on household disposable income. The latter is a more comprehensive measure that may correlate strongly with longevity. This is especially the case for women, whose individual income often diverges from household income. Third, our study ranks individuals based on a fixed income measure over

Figure 2: Life expectancy at age 55 by income group and gender

(a) Men



(b) Women

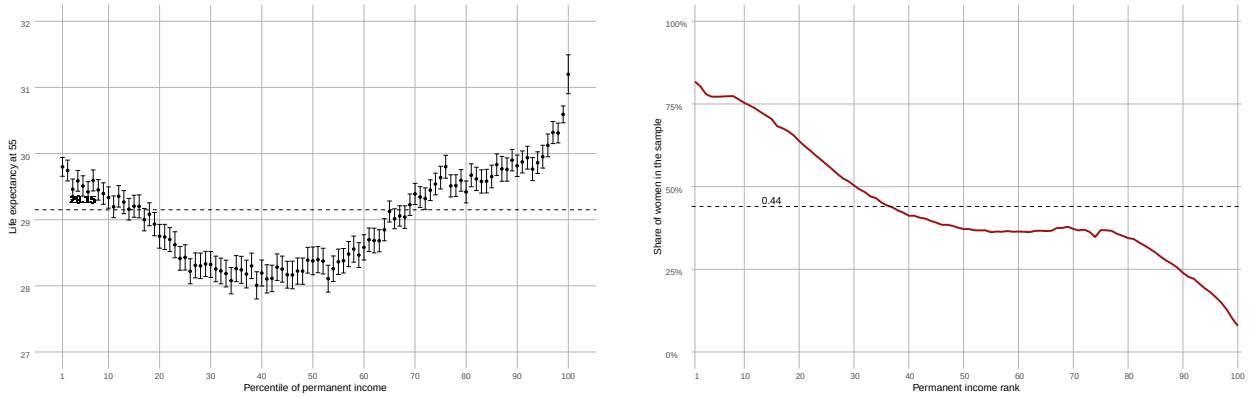


Note: This figure presents life expectancy at 55 by percentile of permanent income – average labour income between 40 and 54. Income rank is defined for the whole sample – men and women pooled. 95% confidence intervals follow the methodology of Camarda (2015). The level of income gradient is estimated by the regression of life expectancy on the income group.
Interpretation: Life expectancy at 55 for a man with median permanent income is around 26 years.
Source: Agirc-Arrco data, authors' calculations

Figure 3: Life expectancy, genders pooled

(a) Life expectancy by income rank

(b) Share of women by income rank



NOTE: This figure presents life expectancy at 55 by percentile of permanent income – average labour income between age 40 and 54, for the whole sample (left). The figure on the right presents share of women by income rank in the sample. Income rank is defined for the whole sample – men and women pooled.
INTERPRETATION: Life expectancy is below average between the 17^{th} and 68^{th} percentiles. There are less women than average above the 37^{th} percentile.
SOURCE: Agirc-Arrco data, authors' calculations

a relatively large age window, thus allowing for income mobility within this time-frame. In contrast, Blanpain (2019) defines income ranks based on the previous year's income. Not accounting for income mobility might overestimate longevity gradients, as shown by Dahl et al. (2021).

4.2 Overall redistribution in the pension system

We have mentioned that pension systems deviate from a purely contributory system – which would provide the same return on contributions for everybody – in three ways.

First, solidarity components of the system aim to ensure sufficient pension levels for individuals with low levels of contribution during their careers. Provided that low-income workers indeed benefit more from these non-contributory components, one expects RC to be decreasing with permanent income, translating into redistribution from high- to low-income workers. Second, depending on career profile, individuals vary in their ability to retire early. The impact on the RC gradient by permanent income is not obvious *a priori*: RC might be lower at the bottom of the gradient – due to more career interruptions and fewer validated periods – as well as at the top – due to lengthy period of studies and late career starts. Third, mortality differentials should lower RC at the bottom of the distribution, all other things being equal.

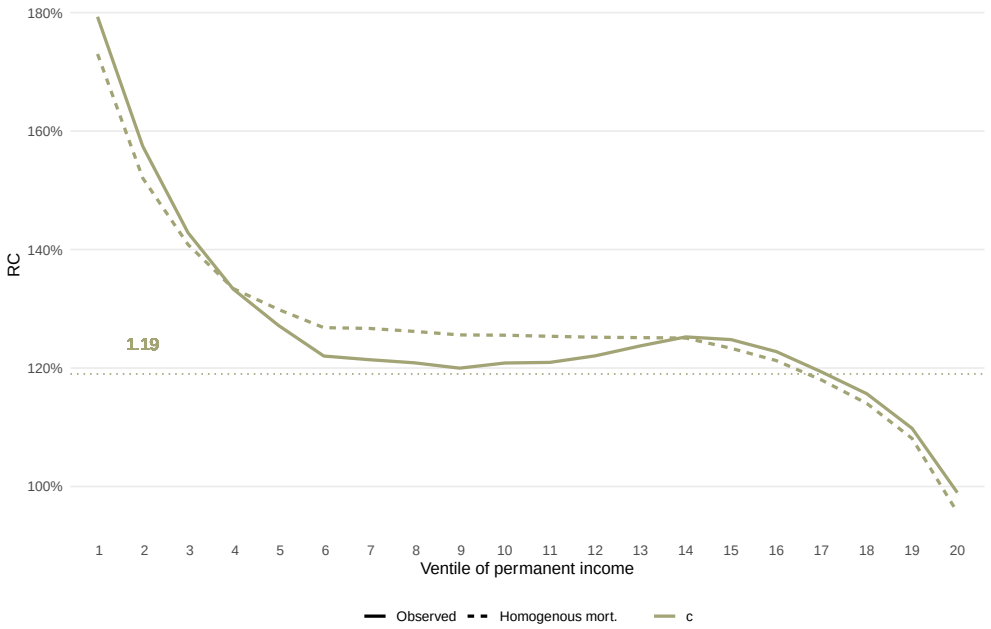
To assess which of these components dominates, we first compute the observed

RC by income group for the full population of our sample, pooling men and women together. The slope of the RC can then be interpreted as the overall progressivity in the French private sector's pension system. This result is presented as the light green line of Figure 4. RC is globally decreasing with income, ranging from 179% for the first ventiles to 99% for the last ones. The overall progressivity of the system is mostly driven by the extreme parts of the income distribution, as RC is relatively flat between the 6^{th} and 15^{th} ventiles.

Mortality differentials do not substantially affect this redistributive pattern. Should the whole population experience the same average mortality from age 55 onwards, the gradient of returns on contribution would be very similar – as shown by the dashed line on the graph. This result can be related to the discussion of Figure 3. Women – with relatively high life expectancy – are overrepresented at the bottom of the lifetime income distribution. Their life expectancy at 55 is close to that of men at the top of the distribution. This composition effect therefore hides some important within- and between-gender differences that do not appear in the overall shape of the RC gradient by quantile presented in Figure 4.

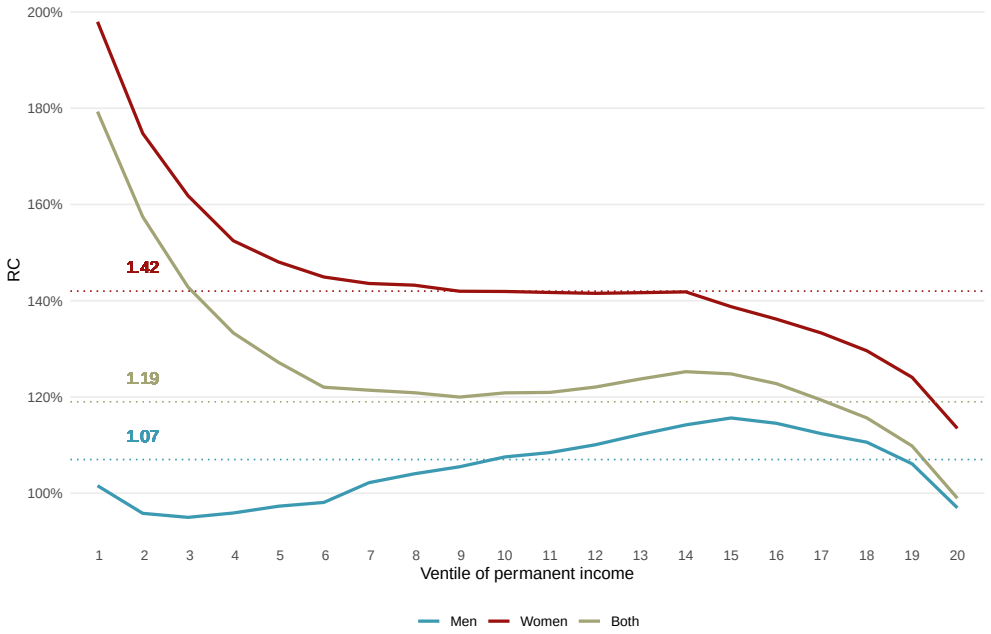
When examining the same gradient by gender in Figure 5, we indeed observe the following patterns: first, RC is higher among women on average (142% against 107%), and for every income group. This suggests that the pension system largely redistributes from men to women. Second, when examining redistribution within gender, we observe opposite patterns among men and women. Among women, the system appears to be highly progressive. Conversely, for men, we find an increasing RC across most of the distribution, suggesting overall regressivity. The RC gradient for both genders pooled, shown in Figure 5, is the aggregation of these results. The bottom ventiles consist of a majority of women: progressivity dominates at the bottom. The top ventiles consist of a majority of men, but we observe progressivity among both men and women: progressivity also dominates at the top. In the middle of the distribution, progressivity among women is more or less offset by regressivity among men.

Figure 4: Return on contributions for cohort 1954



NOTE: This figure presents RC by ventile of permanent income – average labour income between 40 and 54 – for the whole population, with observed (solid line) and homogeneous mortality (dashed line). Income rank is defined for the whole sample – men and women pooled. Observed RC is computed using observed contributions, pension levels, retirement age, and estimated mortality at the income ventile × gender level. It is then aggregated at the income ventile level. Dashed line considers counterfactual returns with homogeneous mortality rates from 55 on for all genders and income ventiles. Discount rate equals 0.98.
INTERPRETATION: Average RC of individuals at the 10^{th} ventile equals 121%. Should they experience average mortality, it would equal 126%.
SOURCE: Agirc-Arrco data, authors' calculations

Figure 5: Return on contributions for cohort 1954



NOTE: This figure presents RC by ventile of permanent income – average labour income between 40 and 54 – for the whole population (green) and by gender (red and blue lines). Income rank is defined for the whole sample – men and women pooled. RC is computed using observed contributions, pension levels, retirement age, and estimated mortality at the income ventile × gender level. It is then aggregated at the income ventile level for the dark green line. Dotted line represents average RC by population – by gender or both genders pooled. Discount rate equals 0.98.
INTERPRETATION: Average RC among women equals 142%. On average, women below the 9^{th} ventile enjoy a higher return than average.
SOURCE: Agirc-Arrco data, authors' calculations

4.3 Between gender redistribution

Figure 6 presents RC disparities observed between men and women (first columns, blue for men, red for women) and their decomposition. We first observe that women experience a 35pp higher return on contributions than men on average. In the absence of disparities in mortality and pension formula between men and women, RC would be 23pp lower for women and 12pp higher for men. This implies a redistribution from men towards women in a lifetime perspective.

We use the decomposition detailed in 3.3 to isolate the role of differentiated mortality from the one of pension formula to this lifetime redistribution.

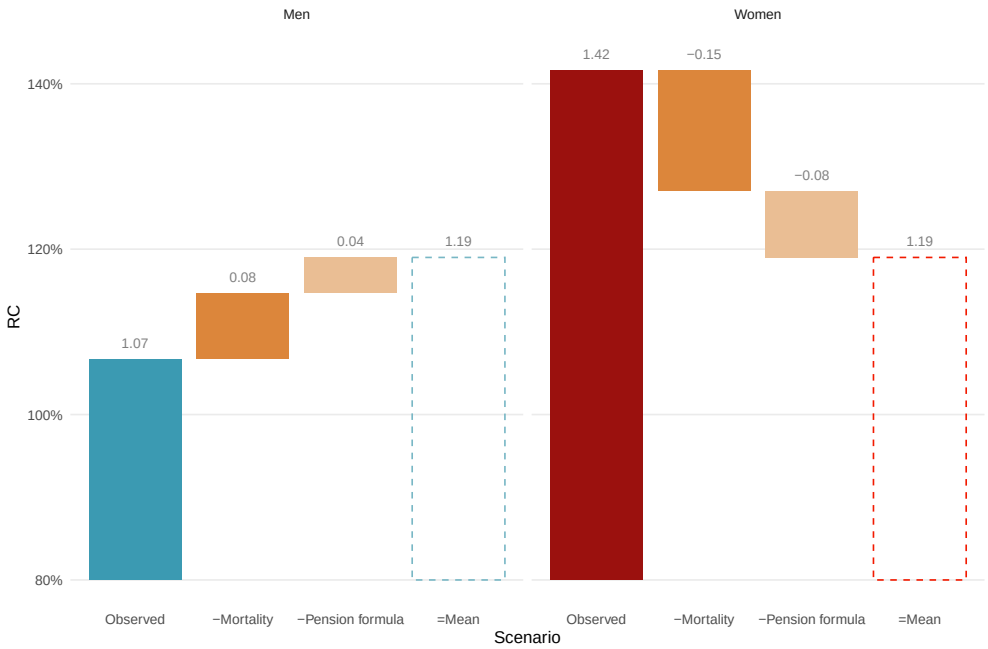
First, differentials in pension formula tend to increase women's RC relative to men's. This corresponds to the rules of the pension formula that alter pension levels relative to lifetime contributions. Part of these rules aim to provide decent pension levels and to partially compensate for career interruptions. Women are indeed more prone to experience such periods of work interruptions. They also benefit more from specific schemes providing additional pension rights for raising children.

In Appendix Figure A.6, we show that at each level of permanent income – mean wage during work periods – women's lifetime contributions are lower than their male counterparts on average. In Appendix Figure A.7, we show that pension rules compensate for this by providing women with lower payback durations (number of years of pensions necessary to cover the amount of lifetime contributions), especially at the bottom of the gradient.

Second, we find that mortality differentials also tend to increase the return disparity: higher longevity among women – refer to Figure 2 – leads to higher returns. It accounts for around two thirds of the gap and has therefore a much stronger impact than the redistribution embedded in the pension formula.

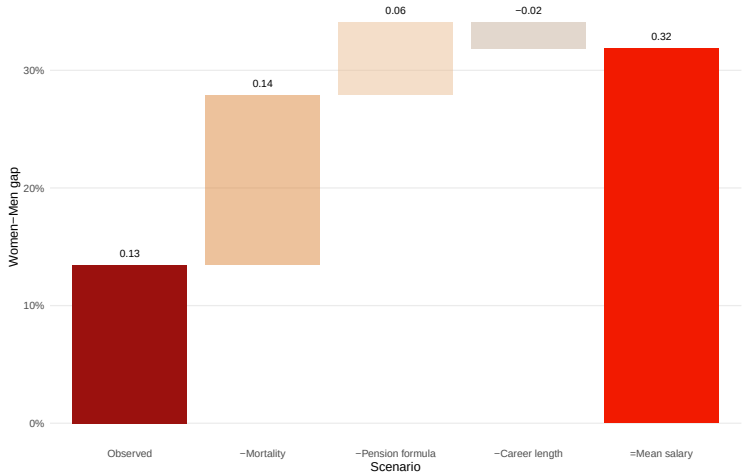
The interpretation of this comparison based on returns on contributions in terms of welfare is not straightforward: women benefit strongly from redistribution on average, but it does not mean that they benefit from higher standards of living at old age. It does not even mean that they benefit from higher levels of lifetime received pensions. We document and decompose this gender gap in lifetime received pensions – Social Security Wealth – in Figure 7. The gender gap in SSW is equal to 13% and is presented in the first column of the figure. Other columns present the contribution of four factors to this observed gap – inequality in mortality, pension formula, career duration and mean salary. We first find that without higher longevity among women, the SSW gap would be 14pp (108%) higher, at 27%. Differential mortality thus reduces the SSW gap by half. Removing also the higher solidarity of the pension formula towards women, the SSW gap would even reach 34%. We also find the effect of mortality to be twice as large as that of the pension formula. The welfare effect is however not the same: explicit solidarity of the pension formula increases the standard of living in each period, while redistribution through mortality corresponds to the realisation of insurance against old age poverty, to which women are more exposed on average. Finally, the remaining gap after removing mortality and pension formula differential is mainly explained by the gap in mean salary during the career –that is, the average gender pay gap for this generation.

Figure 6: Decomposition of the RC gap between gender



Note: This figure presents average RC by gender and the decomposition of the gap between men and women. In the first column, we depict observed return on contributions by gender. Two next columns successively remove differentials in mortality and solidarity resulting from the pension formula between men and women, and depict the marginal effect on RC by gender. Finally, last column amounts to RC without any differential between men and women and amounts to average RC in the overall cohort. Discount rate equals 0.98.
Interpretation: The average observed return is higher among women. Both differentials in benefit from solidarity of the pension formula, and in mortality, contribute positively to this result.
Source: Agirc-Arrco data, authors' calculations

Figure 7: Decomposition of the lifetime pension gap between gender



NOTE: This figure presents the observed lifetime pension gap between men and women and its decomposition into four components. In the first column, we depict observed SSW gap. Four next columns successively remove differentials in mortality, solidarity from the pension formula, career length between men and women, and depict the marginal effect on SSW gap. Remaining SSW gap amounts to the gap in mean salary during the career, defined as the gap in permanent labour income. Discount rate equals 0.98.
INTERPRETATION: The average lifetime pension gap between men and women equals 13%. Inequality in mean wages and career length are the main contributors to this observed gap. Solidarity of the pension formula and higher longevity of women significantly reduce the gap. Removing them leads to a higher gap of around 35%.
SOURCE: Agirc-Arrco data, authors' calculations

4.4 Within-gender redistribution

We now decompose the RC profiles observed in Figure 5, within gender and between income groups. The results are presented in Figure 8. For each gender (men in blue and women in red) and ventile of permanent income (represented on the x-axis), the solid lines represent the observed returns on contributions. These observed returns can be compared to the dashed lines representing the scenario where life expectancy differential within gender are neutralised. Average returns on contributions by gender are finally represented by the horizontal dotted lines.

The decompositions provide very different explanations for the observed profiles for men and women. First, for both men and women, the effect of pension formula (observed from the difference between the horizontal dotted line and the dashed lines) has a strong progressive impact on the RC profile – people at the bottom of the distribution benefit more from non-contributivity in the pension formula than people at the top. The effect is, however, far stronger among women at the bottom of the gradient: all other things being equal, RC of women at the 1st ventile is on average 63pp higher because of these non-contributory components – 19pp for men of the 1st ventile.

More importantly, we find that mortality differentials (observed from the differ-

ence between the solid and dashed lines) have regressive effects on the RC profile for both men and women. The magnitude of this effect is, however, very different and relates to the life expectancy gap observed in Figure 2: mortality differentials within men are large, and their effect on the RC profile more than offsets the progressive effect of solidarity across most of the gradient. The regressive profile of RC initially observed among men is thus mainly explained by the large mortality gradient. This result can be compared to that of Haan et al. (2020). After accounting for mortality differentials, both French and German pension systems appear globally regressive amongst men. Both systems retain some progressivity at the very bottom of the income distribution, reflecting the importance of solidarity components in the pension formula. Results differ at the top of the distribution, where the German system appears to be even more regressive, while the French system shows some progressivity.

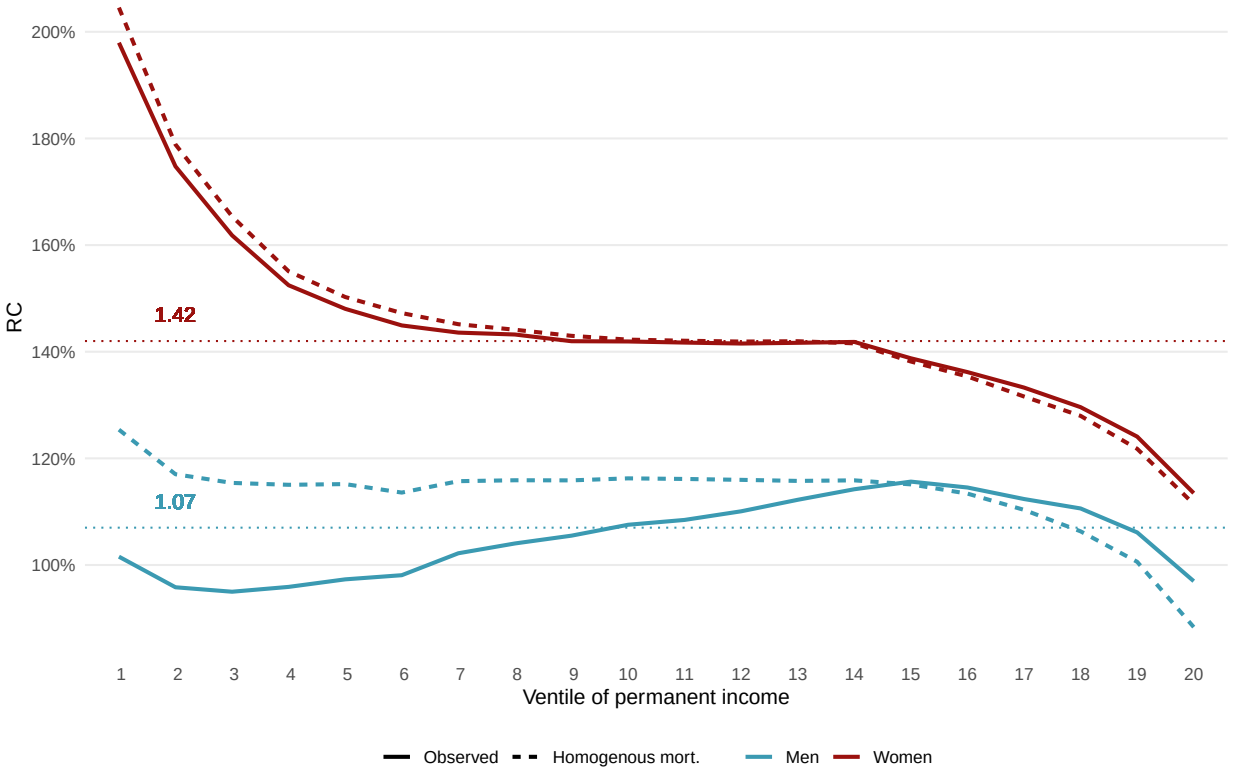
Among women, the effect of mortality differentials we estimate are moderate and the magnitude of the effect on progressivity is very low compared to that of the pension formula. The resulting profiles are therefore different: among men, regressivity dominates between the 3rd and 15th ventile, while progressivity dominates along the entire distribution among women.

5 Conclusion

This paper documents inequalities in life expectancy at 55 between genders and across levels of permanent income among French workers of the private sector. We demonstrate that these inequalities imply large hidden lifetime redistribution in lifetime earnings within the pension system, both between and within gender. In terms of between gender inequality, we find that women experience a higher return on their lifetime contributions than men. This result is not mostly driven by the explicit redistributive features of the pension system, but rather by the average 5.7 life expectancy gap between women and men. However, the gender pension gap remains sizeable despite this redistribution, as the observed amount of pensions received during lifetime is still 13% smaller for women.

Within gender, we find significant regressive impact of mortality differentials across levels of permanent income. This result is particularly salient among men: the 7.2-year differential in life expectancy between the two extremes of the income distribution more than offsets the explicit progressivity of the pension formula. Within women, the life expectancy gradient we observe is smaller, at around 1.5

Figure 8: Return on contributions for cohort 1954, effect of mortality differentials within gender



NOTE: This figure presents RC by ventile of permanent income – average labour income between 40 and 54 – for women (red) and men (blue), with observed (solid line) and homogeneous mortality (dashed line). Income rank is defined for the whole sample – men and women pooled. Observed RC is computed using observed contributions, pension levels, retirement age, and estimated mortality at the income ventile × gender level. Dashed line considers counterfactual returns with mean mortality rates within each gender from 55 on. Dotted line depicts average RC within gender. Discount rate equals 0.98.
INTERPRETATION: The average return for men of the 5^{th} ventile is lower than average: should all men experience average mortality, it would be above average.
SOURCE: Agirc-Arrco data, authors' calculations

years, and its observed redistributive impact is therefore also less prominent.

The implication of this hidden redistribution for policy design is not straightforward (Bozio, Rabaté, & Tô, 2024). A first difficulty is that we may want to compensate for one regressive dimension – between income groups – and not the progressive one – between gender, which can be challenging in practice. A second and more conceptual question is whether this hidden redistribution should even be tackled at all. On one hand, lifetime redistribution is one of the goal and purpose of public pension systems (Diamond, 2004, Bommier et al., 2011). On the other hand, insurance against longevity – which is arguably the main purpose of any pension system – requires some heterogeneity in the realisation of the mortality risk (Liebman, 2002). In that sense, trying to create a negative correlation between effective

replacement rates and predicted longevity is maybe unjustified. All the more so when the risk of mistagging is important (Baurin, 2021, Börsch-Supan et al., 2022, Aubert, 2025): even when based on granular income gradients, within-group variation in life expectancy is likely to remain high. More empirical and normative analyses are therefore warranted to assess if and how differential mortality should be compensated for within the pension system.

Another fruitful avenue for future research is to pursue and extend our analysis of the effect of differential longevity on the overall inequality and redistribution between gender and income groups. Most studies of income inequality only consider instantaneous inequality (Piketty et al., 2018, Bozio, Garbinti, et al., 2024a). This ignores redistribution happening through inequality in longevity: richer individuals contribute more, and longer, to the tax system, but also benefit longer, especially from health-related transfers at old age. Whether the second effect dominates the other – that is whether the tax and transfer system is redistributive on a lifetime basis – is still an open question to this date, despite potential large impacts on equity, efficiency, but also political acceptability of the tax system.

References

Aaberge, R., & Mogstad, M. (2015). Inequality in current and lifetime income. *Social Choice and Welfare*, *44*(2), 217–230.

Attanasio, O. P., & Hoynes, H. W. (2000). Differential mortality and wealth accumulation.

Aubert, P. (2012). Les âges de sortie d'activité. *Revue française des affaires sociales*(4), 79-83.

Aubert, P. (2025). Allowing those with shorter life expectancy to retire earlier: A review of reforms since the 1970s. *Economie et Statistique / Economics and Statistics*. (forthcoming)

Aubert, P., Andrieux, V., & Chantel, C. (2012). Montant des pensions de retraite et taux de remplacement. *Solidarité santé. Etudes statistiques*(33).

Aubert, P., & Bachelet, M. (2012). Disparités de montant de pension et redistribution dans le système de retraite français. *INSEE, série des documents de travail de la Direction des Études et Synthèses Économiques, G, 6*.

Baurin, A. (2021). The limited power of socioeconomic status to predict lifespan: Implications for pension policy. *Journal of the Economics of Ageing*, *20*(1).

Bertrand, M. (2011). New perspectives on gender. In O. Ashenfelter & D. Card (Eds.), (1st ed., Vol. 4B, p. 1543-1590). Elsevier. Retrieved from https://EconPapers.repec.org/RePEc:eee:labchp:5-17

Bertrand, M. (2020, May). Gender in the twenty-first century. *AEA Papers and Proceedings*, *110*, 1-24. Retrieved from https://www.aeaweb.org/articles?id=10.1257/pandp.20201126 doi: 10.1257/pandp.20201126

Betti, G., Bettio, F., Georgiadis, T., & Tinios, P. (2015). *Unequal ageing in europe: Women's independence and pensions*. Springer.

Blanpain, N. (2019, June). *L'espérance de vie par niveau de vie - méthode et principaux résultats, n° F1801.* Retrieved from https://hal.archives-ouvertes.fr/hal-02150680 (working paper or preprint)

Blau, F. D., & Kahn, L. M. (2017, September). The gender wage gap: Extent, trends, and explanations. *Journal of Economic Literature*, *55*(3), 789-865. Retrieved from https://www.aeaweb.org/articles?id=10.1257/jel.20160995 doi: 10.1257/jel.20160995

Bommier, A., Leroux, M.-L., & Lozachmeur, J.-M. (2011). Differential mortality and social security. *The Canadian Journal of Economics*, *44*(1).

Börsch-Supan, A., Bucher-Koenen, T., Goll, N., & Hanemann, F. (2022). Targets missed: three case studies exploiting the linked share-rv data. *Journal of Pension Economics & Finance*, *21*(1), 1–21.

Bozio, A., Garbinti, B., Goupille-Lebret, J., Guillot, M., & Piketty, T. (2024a). Predistribution versus redistribution: Evidence from france and the united states. *American Economic Journal: Applied Economics*, *16*(2), 31–65.

Bozio, A., Garbinti, B., Goupille-Lebret, J., Guillot, M., & Piketty, T. (2024b, April). Predistribution versus redistribution: Evidence from france and the united states. *American Economic Journal: Applied Economics*, *16*(2), 31-65. doi: 10.1257/app.20220023

Bozio, A., Rabaté, S., & Tô, M. (2024). Inequality in life expectancy and the design of pension systems. *Rethinking Pension Reform*.

Bruil, A., Van Essen, C., Leenders, W., Lejour, A., Möhlmann, J., & Rabaté, S. (2022). Inequality and redistribution in the netherlands.

Camarda, C. G. (2015). *Life expectancy confidence interval.* (https://sites.google.com/site/carlogiovannicamarda/r-stuff/life-expectancy-confidence-interval)

Chetty, R., Stepner, M., Abraham, S., Lin, S., Scuderi, B., Turner, N., ... Cutler, D. (2016, 04). The Association Between Income and Life Expectancy in the United States, 2001-2014. *JAMA*, *315*(16), 1750-1766. Retrieved from https://doi.org/10.1001/jama.2016.4226 doi: 10.1001/jama.2016.4226

COR. (2009, January). *L'évolution des paramètres des régimes arrco et agirc.* Retrieved from https://www.cor-retraites.fr/sites/default/files/2019-06/doc-1073.pdf (working paper or preprint)

Dahl, G. B., Kreiner, C. T., Nielsen, T. H., & Serena, B. L. (2021, 12). Understanding the Rise in Life Expectancy Inequality. *The Review of Economics and Statistics*, 1-33. Retrieved from https://doi.org/10.1162/rest_a_01148 doi: 10.1162/rest_a_01148

Diamond, P. (2004, March). Social security. *American Economic Review*, *94*(1), 1-24. Retrieved from https://www.aeaweb.org/articles?id=10.1257/000282804322970670 doi: 10.1257/000282804322970670

DREES. (2023). *Les retraites et les retraités 2023.* (https://drees.solidarites-sante.gouv.fr/sites/default/files/2023-07/RETRAITES23MAJ1707.pdf)

Dubois, Y., & Marino, A. (2015). Le taux de rendement interne du système de retraite français: quelle redistribution au sein d'une génération et quelle évolution entre générations? *Economie et statistique*, *481*(1), 77–95.

Goldin, C. (2014). A grand gender convergence: Its last chapter. *American economic review*, *104*(4), 1091–1119.

Goldin, C., & Lleras-Muney, A. (2019). $XX > XY$?: The changing female advantage in life expectancy. *Journal of Health Economics*.

Guvenen, F., Kaplan, G., Song, J., & Weidner, J. (2022). Lifetime earnings in the united states over six decades. *American Economic Journal: Applied Economics*, *14*(4), 446–479.

Guvenen, F., Karahan, F., Ozkan, S., & Song, J. (2021). What do data on millions of us workers reveal about lifecycle earnings dynamics? *Econometrica*, *89*(5), 2303–2339.

Haan, P., Kemptner, D., & Lüthen, H. (2020). The rising longevity gap by lifetime earnings – distributional implications for the pension system. *The Journal of the Economics of Ageing*, *17*, 100199. Retrieved from https://www.sciencedirect.com/science/article/pii/S2212828X18301452 doi: https://doi.org/10.1016/j.jeoa.2019.100199

Hagen, J., Laun, L., Lucke, C., & Palme, M. (2025). The rising income gradient in life expectancy in sweden over six decades. *Proceedings of the National Academy of Sciences*, *122*(14), e2418145122.

Insee. (2023a). *Purchasing power index of the average annual net salary for full-time jobs in the private sector - men.* Retrieved 2023-04-04, from https://www.insee.fr/en/statistiques/serie/010752373 (Series ID: 010752373)

Insee. (2023b). *Purchasing power index of the average annual net salary for full-time jobs in the private sector - women.* Retrieved 2023-04-04, from https://www.insee.fr/en/statistiques/serie/010752374 (Series ID: 010752374)

Ipp tax and benefit tables. (2023). https://www.ipp.eu/en/ipp-tax-and-benefit-tables/social-security-contributions/. (Accessed: 2023-04-04)

Kinge, J. M., Modalsli, J. H., Øverland, S., Gjessing, H. K., Tollånes, M. C., Knudsen, A. K., ... Vollset, S. E. (2019, 05). Association of Household Income With Life Expectancy and Cause-Specific Mortality in Norway, 2005-2015. *JAMA*, *321*(19), 1916-1925. Retrieved from https://doi.org/10.1001/jama.2019.4329 doi: 10.1001/jama.2019.4329

Kleven, H., Landais, C., & Søgaard, J. E. (2019). Children and gender inequality: Evidence from denmark. *American Economic Journal: Applied Economics*, *11*(4), 181–209.

Kopczuk, W., Saez, E., & Song, J. (2010). Earnings inequality and mobility in the united states: Evidence from social security data since 1937. *The Quarterly Journal of Economics*, *125*(1), 91–128.

Kreiner, C. T., Nielsen, T. H., & Serena, B. L. (2018). Role of income mobility for the measurement of inequality in life expectancy. *Proceedings of the National Academy of Sciences*, *115*(46), 11754–11759.

Li, J., & O'Donoghue, C. (2013). A survey of dynamic microsimulation models: uses, model structure amd methodology. *International Journal of microsimulation*, *6*(2), 3–55.

Liebman, J. B. (2002). Redistribution in the current us social security system. In *The distributional aspects of social security and social security reform* (pp. 11–48). University of Chicago Press.

Mazzaferro, C., Morciano, M., & Savegnago, M. (2012). Differential mortality and redistribution in the italian notional defined contribution system. *Journal of Pension Economics & Finance*, *11*(4), 500–530.

Milligan, K., & Schirle, T. (2021). The evolution of longevity: Evidence from canada. *Canadian Journal of Economics*, *54(1)*, 164-192. doi: https://doi.org/10.1111/caje.12497

Neelakantan, U., & Chang, Y. (2010). Gender differences in wealth at retirement. *The American Economic Review*, *100*(2), 362–367. Retrieved 2023-07-19, from http://www.jstor.org/stable/27805020

Piketty, T., Saez, E., & Zucman, G. (2018). Distributional national accounts: methods and estimates for the united states. *The Quarterly Journal of Economics*, *133*(2), 553–609.

Samek Lodovici, M., Drufuca, S., Patrizio, M., & Pesce, F. (2016, July). *The gender pension gap: Differences between mothers and women without children.* Retrieved from https://www.europarl.europa.eu/thinktank/en/document/IPOL_STU(2016)571363 (Study, European Parliament's Policy Department for Citizens' Rights and Constitutional Affairs)

Samuelson, P. A. (1958). An exact consumption-loan model of interest with or without the social contrivance of money. *Journal of Political Economy*, *66*(6), 467–482.

Santos Silva, J. M. C., & Tenreyro, S. (2006). The log of gravity. *The Review of Economics and Statistics*.

Shi, J., & Kolk, M. (2022). How does mortality contribute to lifetime pension inequality? evidence from five decades of swedish taxation data. *Demography*, *59*(5), 1843–1871.

Tréguier, J. (2021). Retirement: Do women and men make the same decisions? *Revue économique*, *72*(6), 881-928. Retrieved from https://www.cairn-int.info/journal-revue-economique-2021-6-page-881.htm doi: 10.3917/reco.726.0881.

Whitehouse, E. (2006). New indicators of 30 oecd countries' pension systems. *Journal of Pension Economics & Finance*, *5*(3), 275–298.

6 Acknowledgment

This research was funded by the Agirc-Arrco, and by the Joint Programming Initiative (JPI) "More years, better lives" (PENSINEQ Project, ANR-21-MYBL-0002).

A Additional Tables and Figures

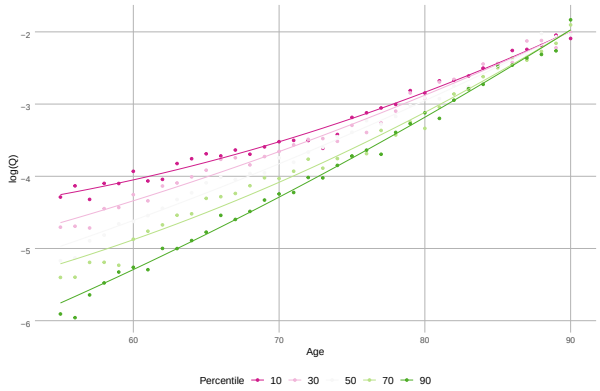
Table A.1: Sample restrictions and sample size

	Restrictions	Gender	Nb of obs. (m)	% previous step	% of Women
	Initial sample	M	14.3	-	
		F	12.3	-	46.2
	Born in France	M	10.8	75	
		F	10.2	83	48.7
	Worked once between 40 and 54	M	6.2	58	
		F	5.8	56	48.1
	>3 years in the AA scheme	M	5.1	82	
		F	4.5	77	46.5
LE sample	Excluding bottom 7%	M	5.0	98	
		F	3.9	88	43.9
	1954 cohort	M	0.21	4.1	
		F	0.20	4.4	45.6
RC sample	Observed pension & full contribution history	M	0.18	87	
		F	0.16	86	45.5

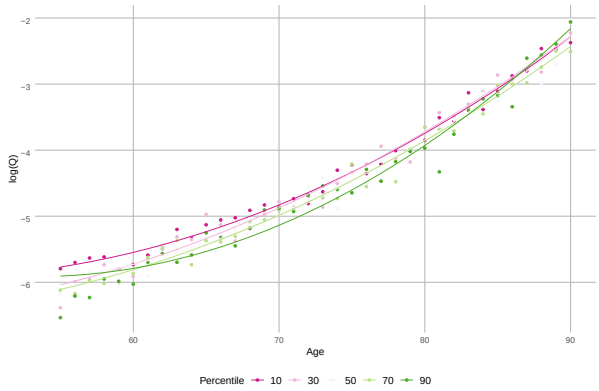
NOTE: This table presents sample size for successive sample restrictions, by gender. The initial sample is made of all *Agirc-Arrco* affiliates for direct rights, of cohorts 1922 to 1964, alive in 2009 and in the year of reaching 55. Each subsequent line represents an additional sample restriction.
SOURCE: Agirc-Arrco data

Figure A.1: Mortality rates by income group

(a) Men



(b) Women

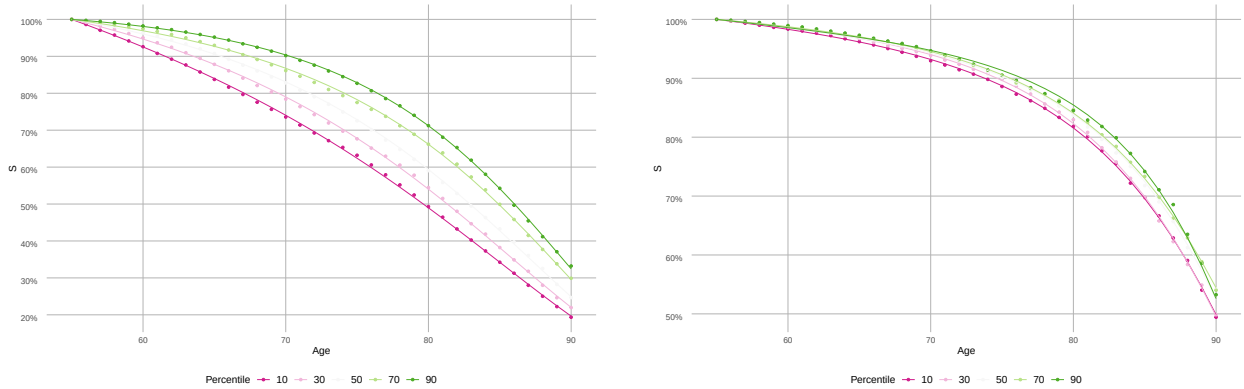


NOTE: This figure presents estimated – line – and observed – dots – log of mortality rates by age for the 10^{th}, 30^{th}, 50^{th}, 70^{th}, 90^{th} percentiles of permanent income, by gender. For each percentile, mortality is estimated out of observed mortality using a Gompertz model between 55 and 90. Income rank is defined for the whole sample – men and women pooled.
INTERPRETATION: A man of the 70^{th} percentile has an estimated mortality rate at 70 of around 1.8% .
SOURCE: Agirc-Arrco data, authors' calculations

Figure A.2: Survival rates by income group

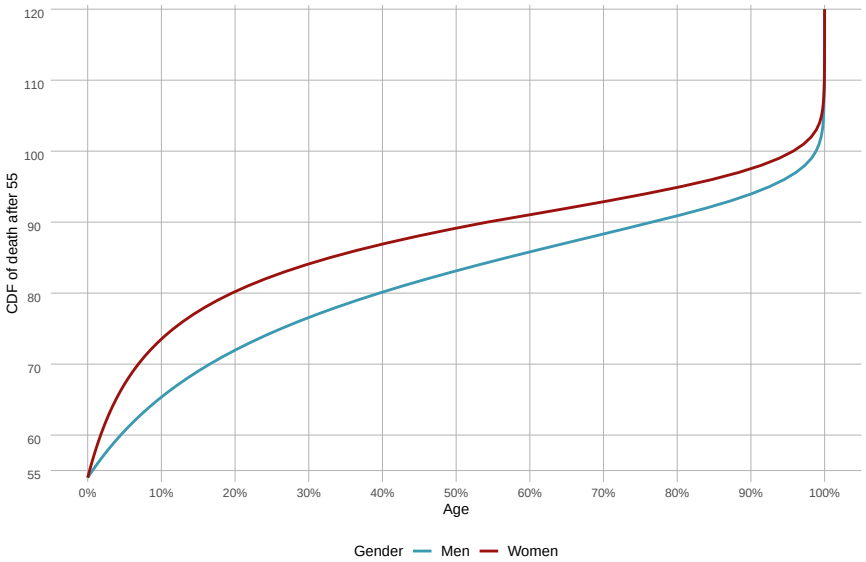
(a) Men

(b) Women



NOTE: This figure presents estimated – line – and observed – dots – survival rates from age 55 by age for the 10^{th}, 30^{th}, 50^{th}, 70^{th}, 90^{th} percentiles of permanent income, by gender. Income rank is defined for the whole sample – men and women pooled. For each percentile, mortality is estimated out of observed mortality using a Gompertz model between 55 and 90.
INTERPRETATION: A man of the 10^{th} percentile alive at 55 has a 50% probability of living up to 80.
SOURCE: Agirc-Arrco data, authors' calculations

Figure A.3: Cumulative distribution function of longevity by gender

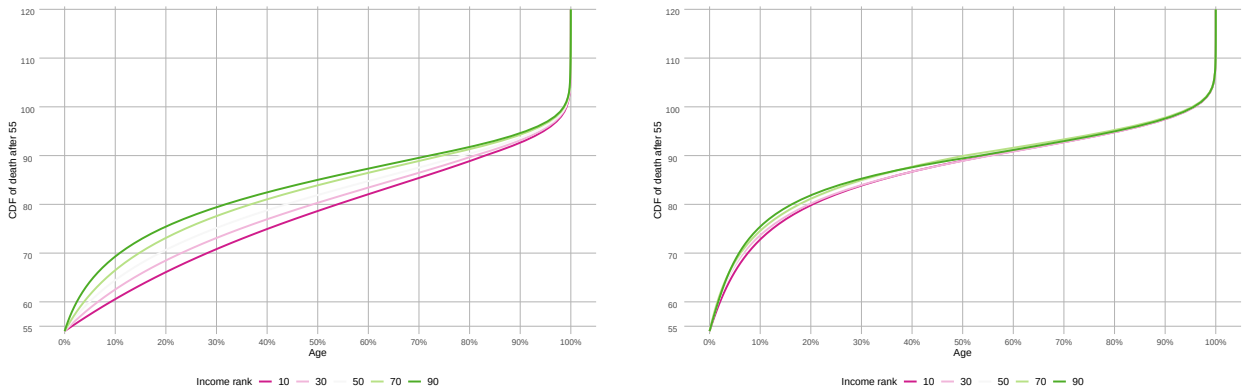


NOTE: This figure presents estimated cumulative distribution function of age of death after 55 by gender. It is computed from estimated mortality rates in section 4.1.
INTERPRETATION: 20% of women alive at 55 die before reaching 80, against 40% of men.
SOURCE: Agirc-Arrco data, authors' calculations

Figure A.4: Cumulative distribution function of longevity by income group and gender

(a) Men

(b) Women

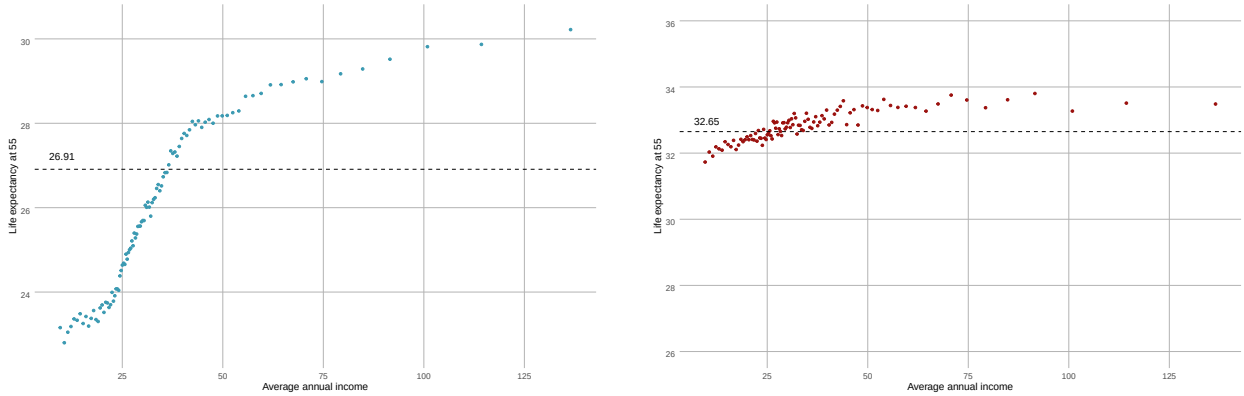


NOTE: This figure presents estimated cumulative distribution function of age of death after 55 by gender and for the 10^{th}, 30^{th}, 50^{th}, 70^{th}, 90^{th} percentiles of permanent income, by gender. Income rank is defined for the whole sample – men and women pooled. It is computed from estimated mortality rates in section 4.1.
INTERPRETATION: Among men of the 10^{th} percentile, 54% die before the age of 80.
SOURCE: Agirc-Arrco data, authors' calculations

Figure A.5: Life expectancy at age 55 by income level

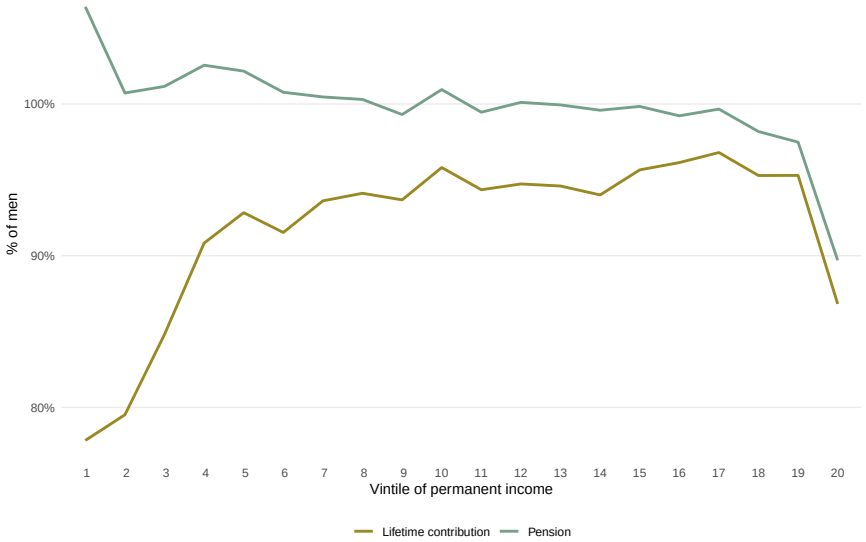
(a) Men

(b) Women



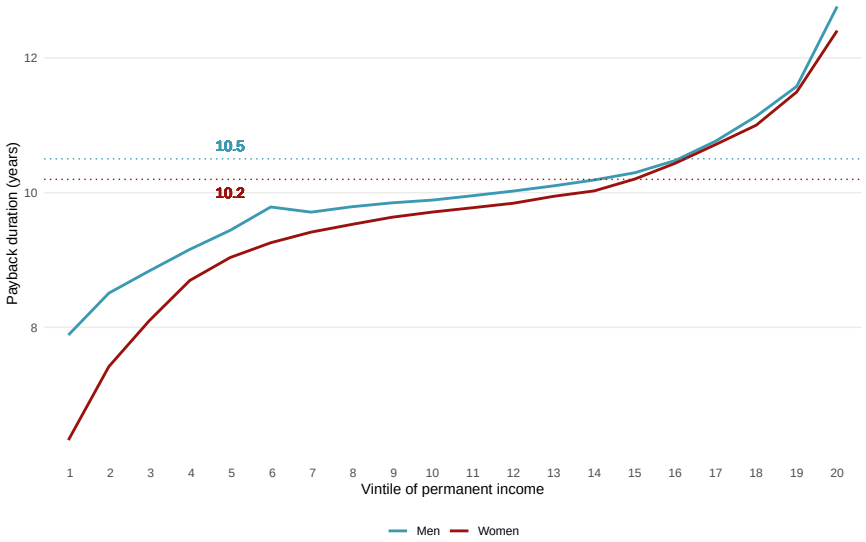
NOTE: This figure presents life expectancy at 55 by level of permanent income – average labour income between 40 and 54 for men. Income rank is defined for the whole sample – men and women pooled. Level of income is expressed in annualised terms and in 2022 €.
INTERPRETATION: Men with average labour income of 50 k€ between 40 and 54 have a life expectancy at 55 of around 28 years.
SOURCE: Agirc-Arrco data, authors' calculations

Figure A.6: Contribution and benefit of women relative to men, by income group



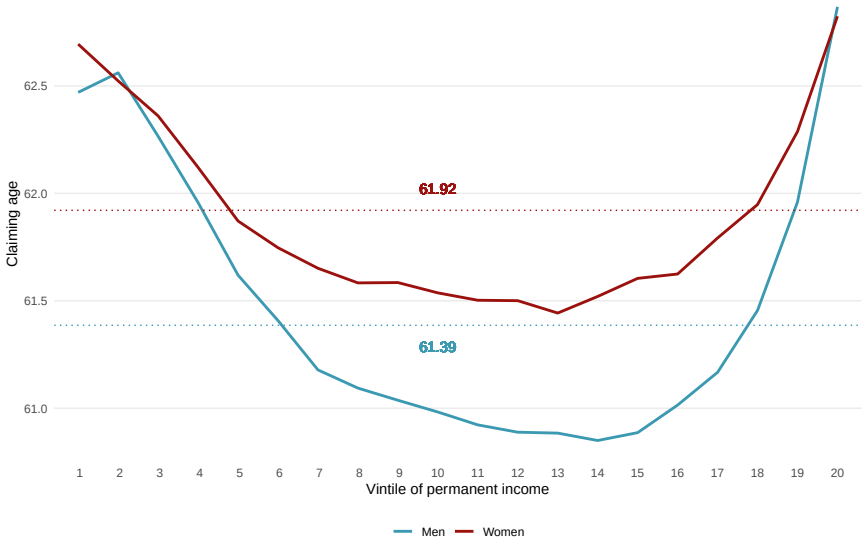
NOTE: This figure presents average pension level and amount of lifetime contributions of women by ventile of labour permanent income, relatively to the same measure among men of the same income group. Income rank is defined for the whole sample – men and women pooled. Discount factor used in computation of lifetime contributions equals 1.
INTERPRETATION: At the 2^{nd} ventile, lifetime pension contribution of women are around 20% lower than men's on average, while pension levels are equal on average. Note that relative pension levels above 100% are observed at the bottom of the gradient, which might seem counterintuitive. It suggests that differential in contributions between men and women is more than compensated by the pension formula, suggesting no contributivity at all. It might be due to some familial rights that affect mostly women.
SOURCE: Agirc-Arrco data, authors' calculations

Figure A.7: Payback duration by gender and income group



NOTE: This figure presents average payback period by ventile of labour permanent income and by gender. Income rank is defined for the whole sample – men and women pooled. Payback period is defined as the ratio between lifetime contributions and pension level at claiming. It corresponds to the pension duration needed to equal lifetime pensions and lifetime contributions. Discount factor used in computation of lifetime contributions equals 1.
INTERPRETATION: On average, men of the 1^{st} ventile need to receive pensions during 8 years for lifetime pensions to equal lifetime contributions – $RC = 100\%$.
SOURCE: Agirc-Arrco data, authors' calculations

Figure A.8: Average claiming age by gender and income group



NOTE: This figure presents average observed claiming age by ventile of labour permanent income and gender. Income rank is defined for the whole sample – men and women pooled. Sample is the same as in section 4 – 1954 cohort. Note that individuals who die before claiming their pension are excluded from the sample.
NOTE: At the 10^{th} ventile, average claiming age of men is 0.6 years lower than retirement age of women.
SOURCE: Agirc-Arrco data, authors' calculations

B Methodological details

B.1 Permanent income: Computation of earnings

Annual earnings are one of the main variables of interest of this study. Data on earnings is however incomplete. However, data on points for earlier periods of the *Agirc-Arrco* records tends to be less incomplete. When earnings are available, we use them. When only points are available, we compute labour income from accumulated points. The last term of equation (A.3) presents a general formula for the relation that exists between annual earnings and the number of points accumulated each year: contributions levied on earnings are converted into points based on the purchase price of the point. In practice, however, the contribution rate is not uniform over total earnings. First, before 2019, the second pillar is split into two regimes, *Agirc* (for executives) and *Arrco* (for non-executives), each with its specific contribution rates and purchase value of points. More precisely, non-executives only contribute to the *Arrco* regime while executives contribute to both *Arrco* and *Agirc* regimes. From 2019, the second pillar scheme and its contribution rules are unified (*Agirc-Arrco*). Secondly, each of these schemes is split into two earnings brackets, with different contribution rates. The threshold between the two brackets is proportional to the Social Security Ceiling (*SSC*, around two times the value of the annual gross minimum wage). Finally, within each bracket, the level of contribution rate might vary from company to company (between legally binding minimum and maximum rates). Therefore, for a given worker i on a given year t, affiliated to a given regime r, earnings are split into two brackets b, each with its specific contribution rates:

$$Point_{irbt} = \frac{W_{irbt}\tau_{irbt}}{PP_{rbt}}$$

Table B.1: Contribution brackets by regime

Regime	Contrib. base	Threshold (SSC)	Non-exec.	Exec.
Arrco	1	<1	yes	yes
	2	1-3	no	yes
Agirc	1	1-4	no	yes
	2	4-8	no	yes
Agirc-Arrco	1	<1	yes	yes
	2	1-8	yes	yes

For this computation, we therefore need to mobilise a set of historical parameters

of the second pillar scheme to recover earnings: contribution rates and purchase values of the points. The *IPP tax and benefit tables* (2023) gathers all French tax and benefit parameters including contribution rates. We should then able to estimate gross wage from the number of contributed points per bracket:

$$W_{irt} = \sum_{b} \frac{Points_{irbt} PP_{rbt}}{\tau_{rbt}} \tag{B.1}$$

This methodology however poses two challenges. First, the inferred earnings are top-coded at the contribution ceiling (for example, three times the Social Security Threshold before 2019 for non-executives). As detailed in Appendix B.2, bunching at the contribution ceiling is not a major issue, as wages above the contribution ceiling are sparse. This suggests that top-coding is not a major issue for the computation of earnings. Secondly, we do not observe actual contribution rates for each individual at each period. We are however able to recover it for periods where both earnings and points are available. On these periods and for each contribution bracket, we define the variable $\Delta_{irbt} = \frac{\tau_{irbt} - \underline{\tau}_{rbt}}{\underline{\tau}_{rbt}}$, where $\underline{\tau}_{rbt}$ denotes minimum legal contribution rate. For each of the four brackets (two brackets per regime), we regress this measure of the distance to minimum rate on a set of fixed effects: individual, year, and quantile of points. We use the result of this simple regression to predict individual rates on missing periods. Note that to predict rates for individuals where no earnings data is observed in the career (only points), we also estimate a similar regression with sex and cohort fixed effects, instead of individual fixed effects. The results of these regressions are found in Table B.2. Figure B.1 displays the distribution of observed and recovered earnings using predictions of the fixed effect models – panel a – and using mean contribution rates applied to contributed points – panel b. Estimated individual contribution rates fit better the overall distribution of wages. This distribution also confirms that top-coding of contribution brackets affects less than 1% of observations.

B.2 Return on contributions: computation of contributions and pensions

On the selected sample, pensions are directly recovered from the *Agirc-Arrco* pension records, which include levels of both first- and second-pillar pensions as well as claiming age for each pillar. More precisely, we observe pension levels from 2014 to 2022. For each individual, we select pension level at the date closest to retirement age. This level is deflated by an index of revaluation of pensions between claiming age and date of observation. There are then deflated by the price index and we

Table B.2: Fixed effect regression – individual contribution rates

Bracket:	Arrco 1		Arrco 2		Agirc 1		Agirc 2	
Fixed-effects								
id	Yes		Yes		Yes		Yes	
year	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
quantile	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
cohort		Yes		Yes		Yes		Yes
sex		Yes		Yes		Yes		Yes
Fit statistics								
Observations	266,935,090	266,935,090	13,864,397	13,864,397	38,394,250	38,394,250	985,148	985,148
Adjusted R^2	0.62416	0.34314	0.74739	0.31929	0.82409	0.68194	0.66702	0.44472

NOTE: *quantile* denotes ventile of earnings, normalised by SMPT and obtained from points using mean contribution rates.

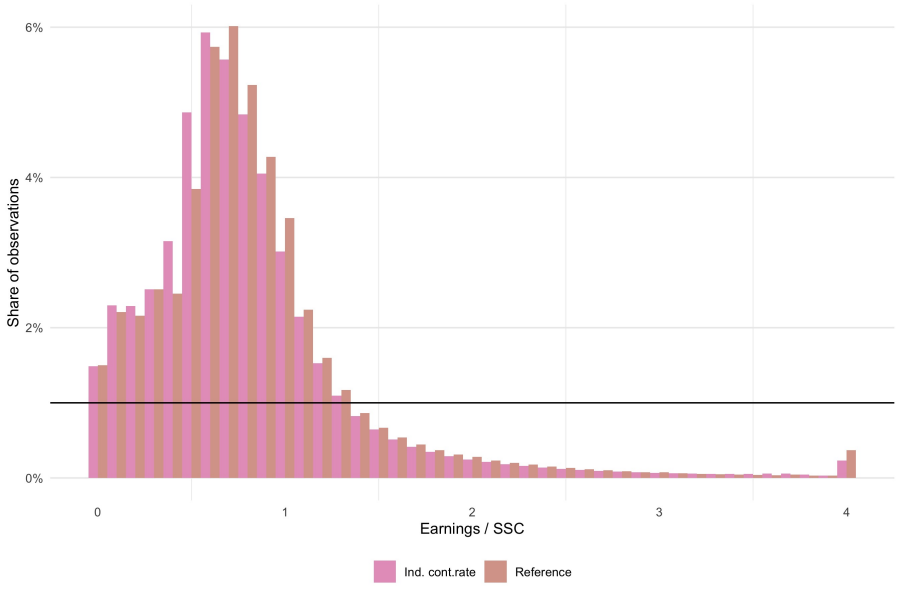
consider that an individual perceives the same amount of pension, in real terms, from claiming age to death. In practice, the evolution of pension levels over years is indeed very close to prices – see Online Appendix A.

Contributions are not directly observed in the *Agirc-Arrco records*, but rather recovered from proxies: salaries and accumulated points observed at each year of the career. We compute contributions in the first pillar from observed salaries and "Cnav" employer and employee contribution rates. In the second pillar, we recover contributions from accumulated points in work periods, the purchase value of each point and the adjustment rate to recover contributions[15]. Observed salaries and points are however not sufficient to recover the whole contribution history in both pillars, for two reasons. First, career databases might be incomplete, especially for the first years of the career, due to data quality, but also because some information on the career are regulated only at claiming – missing work periods in particular. These missing periods are likely to be absent of the career database, but are present as contributed quarters per year in the pensioners' database. Second, the *Agirc-Arrco* career databases only consider periods spent in this second pillar regime. Yet, the "Cnav" first pillar and "Agirc-Arrco" second pillar do not fully overlap: some occupations are affiliated to the *CNAV* scheme and have another second-pillar scheme than *Agirc-Arrco*, and reversely. Although we restrict our analysis to individuals who contributed to both schemes at some point in their career, some of them spent part of their career in these cases. Figure B.2 displays the average share of the career that is missed in the data, in the first and second pillars respectively. As expected, this share is lower in the first pillar, that is directly covered by our data. The share is also rather homogeneous across ventiles and genders, except for

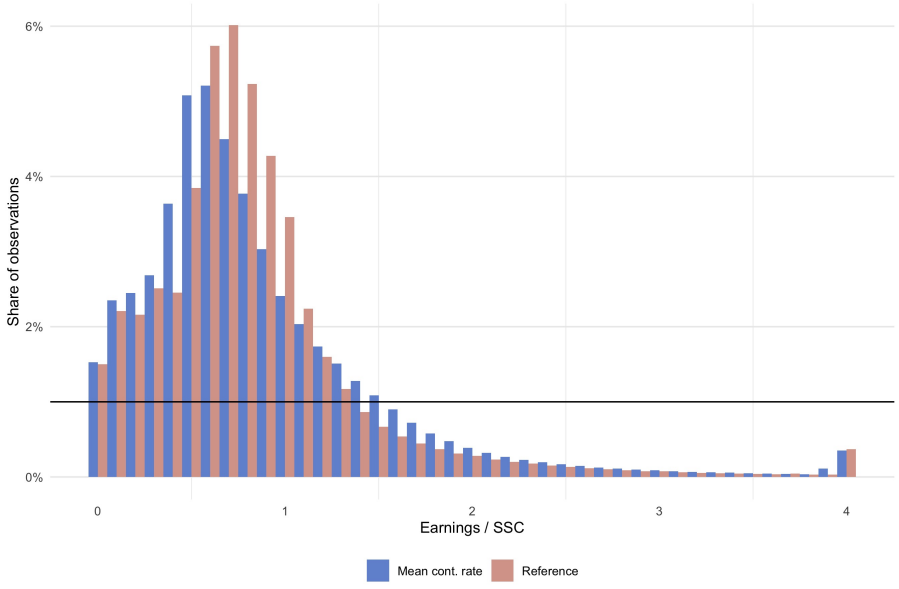
[15]Adjustment rate – "Taux d'appel" – is a factor applied to contribution rates of the second pillar, that increases effective levels of contributions without giving access to pension rights – points

Figure B.1: Distribution of recovered for the 1954 cohort

(a) Estimated individual contribution rates



(b) Mean contribution rates

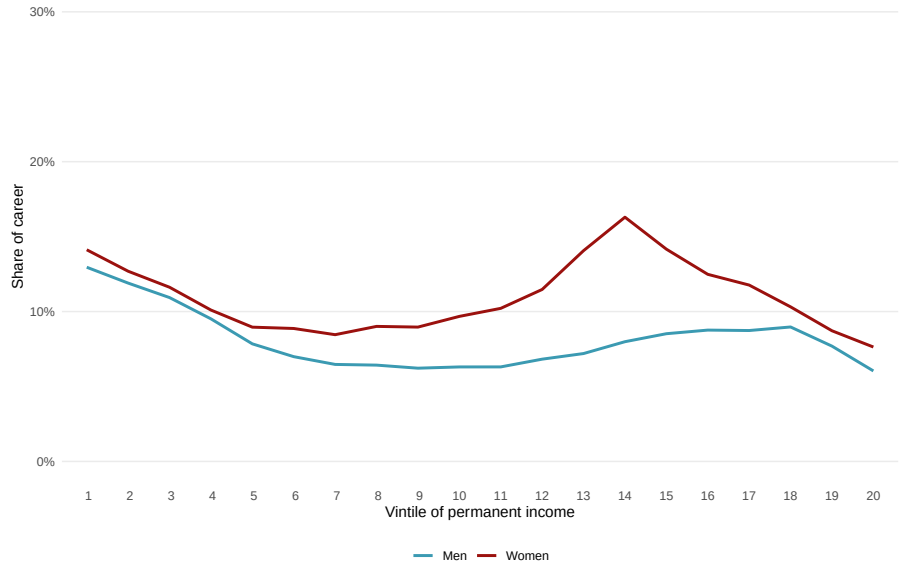


Note: This figure presents the distribution of observed and recovered earnings, normalised by the value of the Social Security Ceiling – SSC and top-coded at 4SSC. Observed earnings are presented in brown. In B.1a, earnings recovered from estimated individual rates are displayed in pink. In B.1b, earnings recovered from estimated individual rates are displayed in blue. The sample is made of one observation by individual-year with positive accumulated points and positive earnings in the career records, for sample born in 1954 used in section 4. Binwidth equals 0.1 SSC.
Interpretation: Around 6% of observed earnings are equal to 0.7 times the SSC.
Source: Agirc-Arrco data

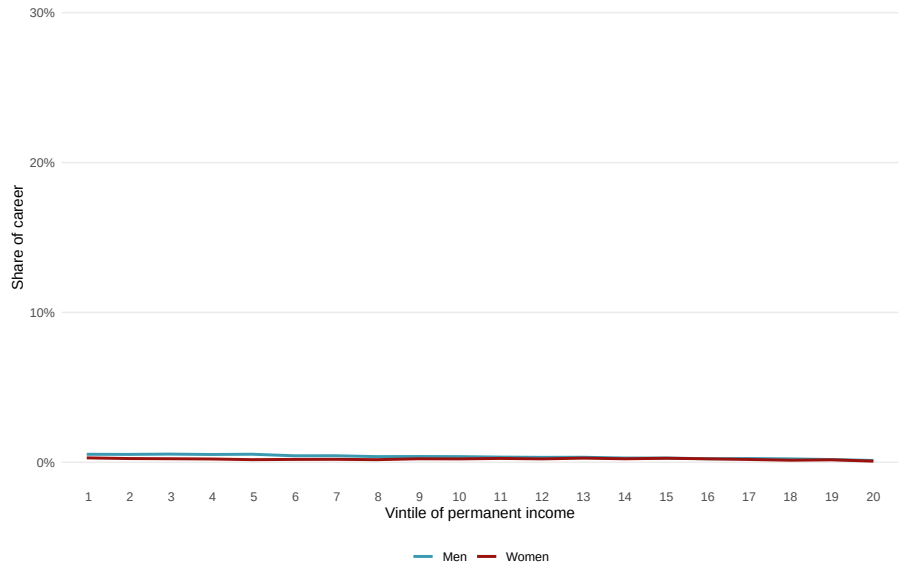
a few exceptions – bottom of the distribution among men and women, upper-middle quantiles among women.

Figure B.2: Mean share of missing career, first and second pillars

(a) First pillar – "Cnav"



(b) Second pillar – "Agirc-Arrco"



NOTE: This figure presents average share of the working career for which contributions are not directly observed by ventile of permanent income and by gender, both for the first and second pillar. Income rank is defined for the whole sample – men and women pooled. Sample is the 1954 cohort used in section 4.
INTERPRETATION: On average, around 10% of *CNAV* work periods among men of the 4^{th} ventile are spent outside of the *Agirc-Arrco* regime or missing in the *Agirc-Arrco* database.
SOURCE: Agirc-Arrco data

To account for these contributions, we estimate salaries on the missing periods in the first and second pillars. We use a simple fixed-effect model on the log of observed salaries using contributed quarters of work and age, fully interacted with permanent income ventile – that do not appear on the formula below for readibility

reasons – and an individual fixed effect as our predictors of earnings:

$$log(salary_{ia}) = \alpha_i + \beta_1 a + \beta_2 a^2 + \beta_3 Qua_{ia} + \beta_4 CumQua_{ia} + \beta_5 CumQua_{ia}^2 + \epsilon_{ia}$$

Where a is age, Qua is worked quarters during the period, $CumQua$ is accumulated quarters as of the beginning of the period, and salary is deflated by an index of mean wages. More precisely, we perform four of these fixed-effect regressions: one per gender and by career period – below 35 and above 35. Salary trajectories indeed appear to be different for each gender, and between the beginning and end of the career. Table B.3 presents the results of these fixed effect regressions. The R^2 levels are quite high in each of these regressions, suggesting a good predicting power of observable variables. We predict this model on periods with observed work periods (contributed quarter) but no observable earnings or points. Contributions are then recovered from these predictions using appropriate contribution rates – and appeal rate in the second pillar. There are a few exceptions: say an individual has only one period of work below 35 and corresponding salary is missing. It is therefore absent of the estimation of the fixed-effect model and we do not predict its salary on the missing period. This is why sample selection excludes a few individuals for whom we do not observe or reconstruct the full contribution history.

Table B.3: Fixed effect regression – log earnings

Dependent Variable:	log(sal)			
Model:	(1)	(2)	(3)	(4)
Variables				
age	0.6503***	-0.0343***	0.4068***	0.0378***
age^2	-0.0103***	0.0003***	-0.0062***	-0.0002***
Qua	0.4679***	0.4356***	0.3609***	0.3737***
$CumQua$	0.0047***	0.0145***	0.0244***	0.0069***
$CumQua^2$	1.25×10^{-5}**	-3.43×10^{-5}***	-0.0002***	-2.56×10^{-5}***
$age \times c$	-0.0005***	0.0016***	0.0017***	0.0004***
$age^2 \times c$	3.78×10^{-5}***	-1.67×10^{-5}***	-5.74×10^{-6}***	-6.53×10^{-6}***
$Qua \times c$	-0.0003***	-0.0025***	-0.0005***	-0.0021***
$CumQua \times c$	0.0002***	-0.0001***	-2.65×10^{-5}	3.52×10^{-5}***
$CumQua^2 \times c$	-5.46×10^{-6}***	5.61×10^{-7}***	-2.41×10^{-6}***	1.33×10^{-7}***
Fixed-effects				
id	Yes	Yes	Yes	Yes
Fit statistics				
R^2	0.78656	0.58752	0.73374	0.58060
Within R^2	0.72982	0.15823	0.58974	0.21794

*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*
(1): Men below 35
(2): Men above 35
(3): Women below 35
(4): Women above 35
NOTE: Qua, $CumQua$, c respectively denote worked quarters in the period, accumulated work quarters as of the beginning of the period and permanent income percentile.

B.3 Return on contribution: Decomposition

We first consider the difference in RC of men and women of the 1954 cohort. We compute them at the population level, relating the total sum of contributions of men (*resp.* women) to the total flow of pensions of the same population. We compare these gender specific RC to the RC of complete 1954 cohort, and we decompose the differences between men (*resp.* women) RC and overall RC to disentangle the importance of the role of differences in retirement age, replacement rate, and life expectancy. In all counterfactual scenarios, we keep constant the overall generosity of the system.

Formally, let Ω be the total population of interest (in our case, the 1954 cohort), and F and M the subpopulation of female and males.

We note RC_F and RC_M the observed return on contributions of female and male, while RC_Ω is the observed RC for the whole 1954 cohort. We then decompose the difference between RC_Ω and the RC_F and RC_M. To do so, we build counterfactual RC using the minimal building blocks that we have in the data, gender $\times$ income ventile. For each block, x, RC is obtained as:

$$RC(P_x, S_x, \pi_x, C_x) = \frac{P_x \sum_{a=55}^{120} \beta^a \pi_{xa} S_{xa}}{\sum_{a=8}^{68} \beta^a C_{xa} S_{xa}} = \frac{SSW(P_x, S_x, \pi_x)}{LTC(S_x, C_x)}$$

where P_x is the sum of pension at retirement for individuals belonging to x (fixed over ages in real terms), $\pi_x = (\pi_{x,55}, \ldots, \pi_{x120})$ is a vector describing the probability of being retired, $S_x = (S_{x8}, \ldots, S_{x120})$ a vector of survival probability, and C_x a vector of contributions by age. We define SSW as social security wealth, and LTC as lifetime contribution. Our main assumption is that we consider retirement age and mortality to be independent and that our sample is supposed to be representative of the whole 1954 cohort conditional on surviving. Note that for any population $K \in \{M, F, \Omega\}$:

$$RC_K = \frac{\sum_{x \in K} \omega_{xK} SSW(P_x, S_x, \pi_x)}{\sum_{x \in K} \omega_{xK} LTC(S_x, C_x)}$$

where ω_{xK} denotes the share of population K made of individuals of building block x, $\sum_{x \in K} \omega_{xK} = 1$. We use the 1954 sample – see Table A.1 to define these shares. Counterfactuals are then built by replacing gender-specific quantities by average ones over the whole population under the constraint that the total RC of the pension system (RC_Ω) remains unchanged.

Life expectancy The first counterfactual accounts for life expectancy and is obtained from the intermediate quantity:

$$\forall \ K \in \{F, M\} : \ \widetilde{RC}_K^{LE} \ = \ \frac{\sum_{x \in K} \omega_{xK} SSW(P_x, S_\Omega, \pi_x)}{\sum_{x \in K} \omega_{xK} LTC(S_\Omega, C_x)}$$

resulting average return is then:

$$RC_\Omega^{LE} = \frac{\sum_{x \in \Omega} \omega_{x\Omega} SSW(P_x, S_\Omega, \pi_x)}{\sum_{x \in \Omega} \omega_{x\Omega} LTC(S_\Omega, C_x)}$$

The counterfactual adjusted for financial equilibrium is thus:

$$RC_K^{LE} \ = \ \widetilde{RC}_K^{LE} \times \frac{RC_\Omega}{RC_\Omega^{LE}}$$

RC_K^{LE} is thus the counterfactual return on contribution of population K if they had the same life expectancy as the whole population Ω, while keeping constant the total amount of pension spending.

Decomposition The decomposition is obtained using the different terms:

$$RC_K - RC_\Omega = \underbrace{RC_K - RC_K^{LE}}_{Impact \ of \ LE} + \underbrace{RC_K^{LE} - RC_\Omega}_{Impact \ of \ pension \ formula}$$

B.4 Social Security Wealth: Decomposition

To decompose the effects of differentials in mortality, replacement rate, retirement age and career on Social Security Wealth between gender, we follow a methodology that is close to the one developed for RC. The population Ω is divided into two populations M and W. We then compute observed and counterfactual SSW for our minimal building block, gender $\times$ income ventile.

For each block x, observed SSW is:

$$SSW(P_x, S_x, \pi_x) = P_x \sum_{a=55}^{120} \beta^a \pi_{xa} S_{xa}$$

where P_x is the sum of pension at retirement for individuals belonging to x (fixed over ages in real terms), π_{xa} describes probablity of being retired, S_{xa} survival

probability. For any population $K \in \{M, F, \Omega\}$:

$$SSW_K = \sum_{x \in K} \omega_{xK} SSW(P_x, S_x, \pi_x)$$

where $\omega_{x,K}$ denotes the share of population K made of individuals of building block x, $\sum_{x \in K} \omega_{xK} = 1$. We use the 1954 sample – see Table A.1 to define these shares. Counterfactuals are built by replacing population-specific quantities by average ones. The main change compared to the RC decomposition comes from the addition of two counterfactuals, to account for differentials in career length and mean salary between men and women. Those differentials in career profile have indeed an impact on SSW but not on RC per say.

The four counterfactuals successively remove differentials in mortality, solidarity from the pension formula, career length, mean salary. SSW is respectively denoted SSW^{LE}, $SSW^{LE,PF}$,,$SSW^{LE,PF,CL}$ and $SSW^{LE,PF,CL,SA}$ in each of these counterfactuals. Formula of the two first counterfactuals are given, for a given population $K \in \{M, F\}$, by:

$$\begin{aligned} SSW_K^{LE} &= \sum_{x \in K} \omega_{xK} SSW(P_x, S_\Omega, \pi_x) \\ SSW_K^{LE,PF} &= \sum_{x \in K} \omega_{xK} SSW(\alpha_\Omega LTC_x, S_\Omega, \pi_\Omega) \end{aligned}$$

where LTC still denotes lifetime contribution and α_Ω denotes the common replacement rate for scenarios without solidarity differential. Note that in this methodology, removing the solidarity from the pension formula is equivalent to providing everyone with the same replacement rate and retirement age. Finally, note that we get $SSW_K^{LE,PF,CL}$ by computing average permanent income within population K. Indeed, one can see that the remaining differential in $SSW_K^{LE,PF}$ is the combination of career length and salary profile. Average permanent income removes the first one.

The SSW gender gap is finally defined by:

$$\Delta SSW = \frac{SSW_M - SSW_W}{SSW_M}$$

We decompose it by noticing that:

$$\begin{aligned}\Delta SSW = & \quad \Delta SSW - \Delta SSW^{LE,PF,CL,SA} \\ = & \quad \underbrace{\Delta SSW - \Delta SSW^{LE}}_{Impact\ of\ LE} + \underbrace{\Delta SSW^{LE} - \Delta SSW^{LE,PF}}_{Impact\ of\ PF} + \\ & \underbrace{\Delta SSW^{LE,PF} - \Delta SSW^{LE,PF,CL}}_{Impact\ of\ CL} + \underbrace{\Delta SSW^{LE,PF,CL} - \Delta SSW^{LE,PF,CL,SA}}_{Impact\ of\ SA}\end{aligned}$$

Note that we do not need to define $SSW^{LE,PF,CL,SA}$ since by definition, $\Delta SSW^{LE,PF,CL,SA} = 0$, *ie* absent any differential in career, retirement and mortality, SSW is the same for men and women.

C Sensitivity analysis

C.1 Permanent income definition

Figure C.1 shows alternative measures of income gradient within and between gender, obtained through variations on the methodology used in section 4.1. Our main estimate of life expectancy gradient appears mostly robust to these variations, that can be classified in six types.

Contribution rate Earnings at the numerator of our permanent income measure are recovered from contributed points and contribution rates. In our main estimate, we use estimated individual rates. Instead, we also try using mean contribution rates by contribution bracket and year, recovered from (COR, 2009) and *IPP tax and benefit tables* (2023) data. We find little impact on the gradient.

Exclusion of incomplete careers We test different restrictions of contribution duration in the *Agirc-Arrco* regime between 40 and 54, from merely one day to 3 years, our main estimate. The higher the threshold, the more individuals with short and non-representative careers in the private sector are excluded. The income gradient tends to increase slightly with this threshold, among men only. This is consistent with what we observed in Online Appendix B: individuals with other sources of individual labour income at the bottom of the gradient are mostly men.

Exclusion of abnormally low earnings We try to exclude observations where observed labour income is lower than minimum wage (with thresholds from 1/4 to 1 minimum wage): daily earnings way below minimum wage might be considered

as "absurd" and coming from data issues. On the other hand, note that they might correspond to part-time labour, which is why we do not exclude them in our main estimate. As could be expected, the higher the restriction, the lower the observed longevity gradient across income groups and between genders.

Inclusion of bottom 7% In our main sample selection, we exclude individuals with the lowest levels of permanent labour income. We indeed know that for a significant part of this subsample, a large share of their career is spent out of the private sector – see Figure B.3. For these, low labour income in the private sector is not suggestive of low total income. However, by doing so, we also exclude some individuals with genuinely low levels of labour income. We therefore try to include this subpopulation as robustness and find that the gradients are mostly left unchanged.

normalisation by career length In our main estimate, we divide the sum of observed private sector earnings by the observed number of years worked during the same period – sum of days work divided by 365. Doing so avoids a downward bias in the rank of individuals with relatively high income who did not spend the whole year working in the private sector. However, it implicitly assumes that earnings outside of work days in the private sector were similar. This might lead to an upward bias in the rank of individuals for whom these period were made of work interruption, which represent negative income shocks. We therefore try normalising labour income by years partially spent in the *Agirc-Arrco regime*, instead of solely days of employment. The effect on the gradient is slightly positive within gender, especially among men.

Age range for permanent income In our main estimate, permanent income is computed on the 40 to 54 age window. The upper bound is defined to avoid reverse causality between income and mortality after 55. The lower bound is a compromise between accounting for income mobility on the whole career (wider age window) and completion of our data source for older cohort and young ages (narrower age window). We test alternative definitions of this lower bound, from 35 to 50. For consistency, we restrict the analysis to the 1927-1964 cohorts (cohorts before 1927 reach the age of 35 before 1962). Variations in the age window may have two opposite effects on the life expectancy gradient. On the one hand, the narrower the age range, the less income mobility is taken into account, which should make the gradient bigger. On the other hand, there is a sample selection effect. People having a working period between 50 and 54 might be relatively wealthier and healthier than

those working between 35 and 54. Figure C.2 seems to corroborate this hypothesis: for a comparable average income level, life expectancy is higher for the 50-54 age window that for the 35-54 age window (respectively in yellow and orange on the graph). Overall, we find little effect of the age window on observed gradient within gender, suggesting that both effects offset each other. The gradient between gender is slightly lower on narrower age windows.

Income rank In our main estimate, we define income thresholds based on percentiles of the distribution for cohort 1964 – whole surviving cohort at 55 observed – and apply them to all cohorts. As robustness, we instead define quantiles of permanent income by cohort. To do so and to account for differential mortality up to the first ages of observation, for older cohorts, we perform the following algorithm: for the 1954-1964 cohorts, we define labour income percentiles by cohort (1% of the population per quantile), and compute observed mortality at 55. It allows to recover the share of each percentile that survives up to 56, and the share of the population observed from 56 on (cohort 1953 in 2009) that enters each percentile. We repeat this process up to cohort 1922. This methodology has the advantage of making no assumption on the evolution of the income distribution across cohorts, provided moratlity was the same in each quantile. The inconvenient is that it relies on observed – and therefore noisy – mortality coefficients of younger cohorts to retropolate income quantiles from older cohorts. We find a small positive effect on the gradient within men.

C.2 Life expectancy computation

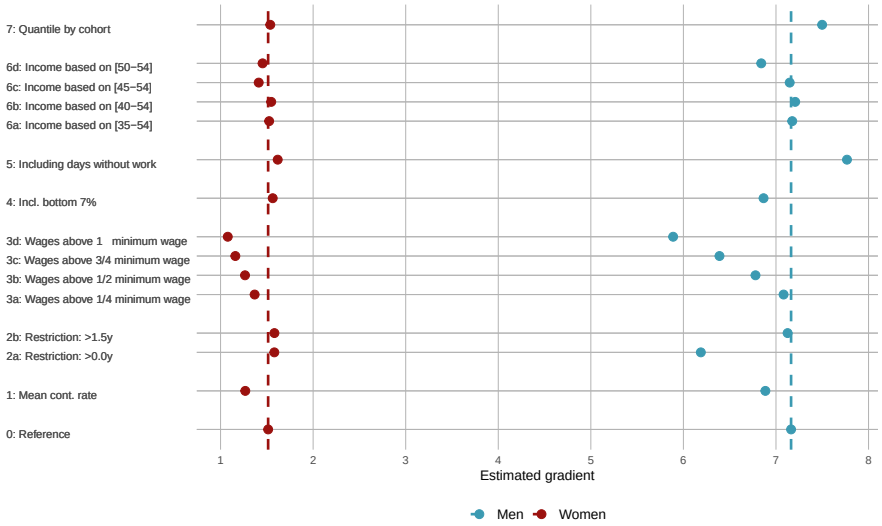
Figure C.3 shows alternative measures of income gradient within and between gender, obtained through variations on the mortality model defined in section 3.2. Our main estimate of life expectancy gradient appears robust to variations in the model's specification. Modifying years or cohorts selected in the sample of the analysis has a bigger impact on the gradient.

Mortality model In our main estimate, we assume log-mortality evolves quadratically with age and estimate the model using Poisson Pseudo-Maximum Likelihood – PPML – between 55 and 90. From 90 to 120, we assume mortality rates equal to the those of the overall population by gender. We try to deviate from this baseline model in four ways.

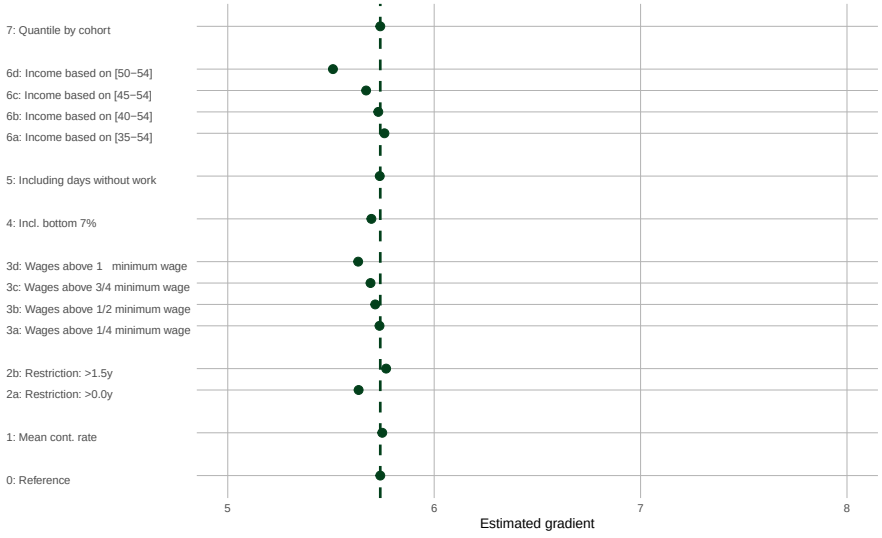
First, we consider a linear specification for the PPML model, instead of a quadratic one:

Figure C.1: Life expectancy by income group: sensitivity analysis

(a) Gradient within gender



(b) Gradient between genders

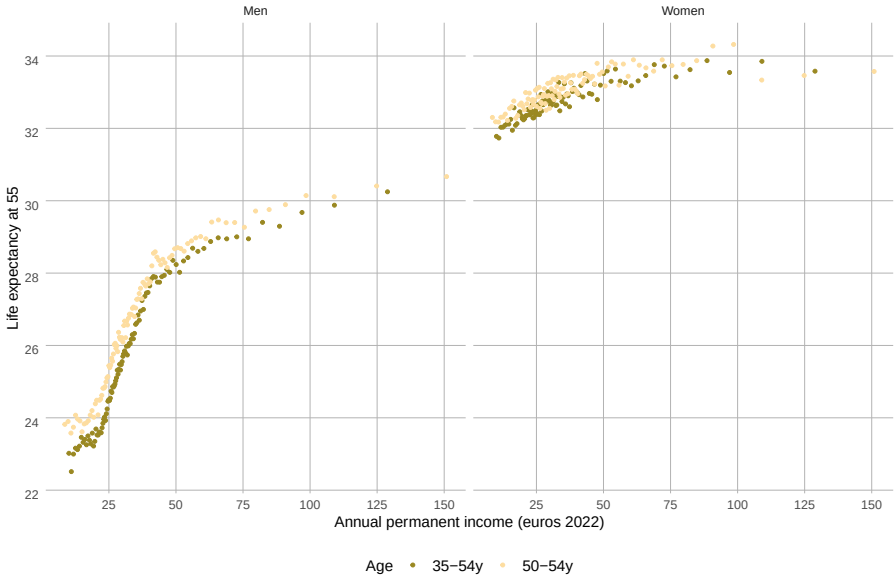


NOTE: This figure presents life expectancy gradients within gender (C.1a) and between genders (C.1b) for different methodologies of permanent income definition and by gender. Life expectancy gradient is defined as the estimated coefficient from the regression of life expectancy at 55 on percentile of labour permanent income. The reference point is the methodology described in 4.1: Average gross labour income for individuals with more than 3 years in the regime between 40 and 54, where permanent income groups are defined for men and women pooled, base on the income distribution of cohort 1964.
INTERPRETATION: Compared to the reference point, the men's gradient when including individuals with at least one work day in the regime is around 1.0 year lower.
SOURCE: Agirc-Arrco data, authors' calculations

$$E(Q_{xa}) = \exp(\alpha_x + \beta_x a) + \varepsilon_{xa}, E(\epsilon_{xa}|a) = 0$$

Second, following Chetty et al. (2016), we assume mortality rates follow a Gom-

Figure C.2: Life expectancy by income level and gender: 35-54 VS 50-54 age window



NOTE: This figure presents life expectancy at 55 by level of permanent income and gender, expressed in annualised terms and in 2022 euros, for two methodologies: average income between 35 and 54 (orange) or 50 and 54 (yellow). In both cases, cohorts are 1927-1964. Income rank is defined for the whole sample – men and women pooled.
INTERPRETATION: For the same average annual income of around 50k € for men, life expectancy is higher in the second sample, averaging income on the 50-54 age window.
SOURCE: Agirc-Arrco data, authors' calculations

pertz model up to 90.

$$\ln\left(Q_{xa}\right) = \alpha_x + \beta_x a + \varepsilon_{xa}$$

Third, we consider a logistic model between 55 and 90. According to the logistic model, the logistic transformation of mortality rates evolves linearly with age for a given sub-population.

$$\ln\left(\frac{Q_{xa}}{1-Q_{xa}}\right) = \alpha_x + \beta_x a + \varepsilon_{xa}$$

Note that we considered a PPML model mainly to deal with a few observations where observed mortality rates are zero. In the Gompertz and logistic model, instead, we deal with them by interpolating these null mortality coefficients from non-null coefficients at adjoining ages. This arbitrary method that can introduce bias is the reason why we prefer PPML specifications in our main estimate. Also note that even absent the zeros, the Gompertz model is not equivalent to the linear PPML specification and needs not result in the same estimates, if residuals are heteroskedastic (from Jensen's inequality, see Santos Silva & Tenreyro (2006)).

Fourth, we consider computing life expectancy out of observed (unsmoothed)

mortality rates.

Overall, these alternative specifications do not have much effect on the gradient within and between gender.

Reweighting observations In our main specification, we put as much weight on each observation (sub-population-age). Implicitly, it puts a lot of weight on few surviving individuals of old cohorts at old ages. We try reweighting observations in the Gompertz model by the number of survivors at the beginning of the year (denominator of the observed mortality rate). The downside of this method is that it puts less weight on relatively old ages at which a large share of deaths of a given cohort is observed. In practice, we observe an almost null impact both within and between genders.

Gradient by year of observation In our main estimate, we pool death and survival counts across eleven years (2009-2019) by subpopulation to estimate our model. It might be that trends in mortality inequality within and between gender made the life expectancy gradient evolve during this period. To verify it, we estimate the same PPML model on each year separately. In 2009, we observe mortality only up to 87 (cohort 1922). For consistency, for each year, we estimate the model using ages 55 to 87, and apply estimated coefficients up to 90. For sample size reason, we do so by ventile of permanent income. We find that between 2009 and 2019, the gradient across income groups increased both among women (+1.2 year) and men (+1.4 years), while the gradient between men and women decreased by 0.7 years. As showed in Figure C.4, during this period, life expectancy increased for whole sub-populations but mostly among men and mostly at the top of the gradient.

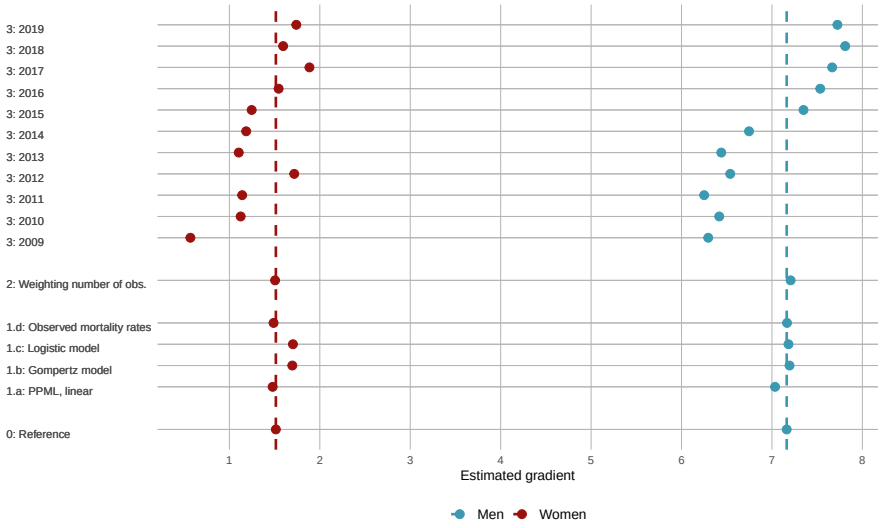
C.3 Computation of RC

Figures C.5 and C.6 respectively display the decomposition of within gender RC gaps, and RC and SSW gap decomposition between gender, using two extreme values of the discount factor: 1 and 0.95. In our main analysis, we used $\beta = 0.98$. A higher β puts more weight on later periods of retirement, where mortality is high and differentials in mortality rate matter more. Differentials in RC and relative contribution of mortality to this gap is therefore stronger with higher levels of β, both within and between gender.

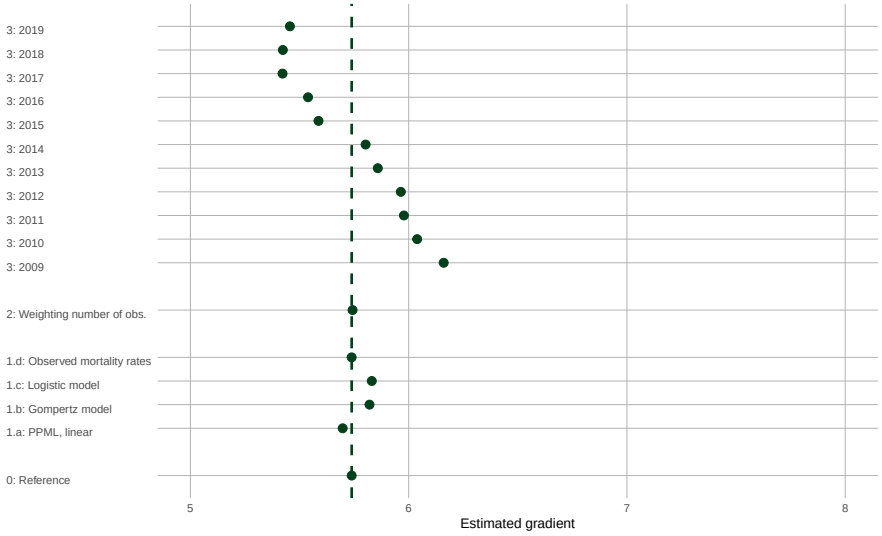
Finally, in Figure C.7, we test the sensitivity of our RC decomposition methodology to the order of decomposition, between gender. We mentioned that the pre-

Figure C.3: Life expectancy gradient: sensitivity to the mortality model

(a) Gradient within gender



(b) Gradient between genders

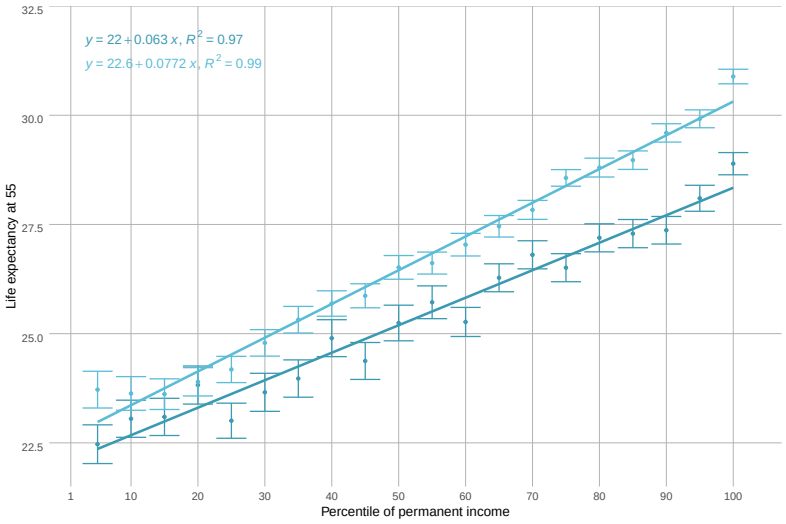


NOTE: This figure presents life expectancy gradients by gender for different mortality models and by gender. Life expectancy gradient is defined as the estimated coefficient from the regression of life expectancy at 55 on percentile of labour permanent income. The reference point is the logistic model described in 4.1.
INTERPRETATION: Compared to the reference point, the women's gradient was around 0.9 year lower in 2009.
SOURCE: Agirc-Arrco data, authors' calculations

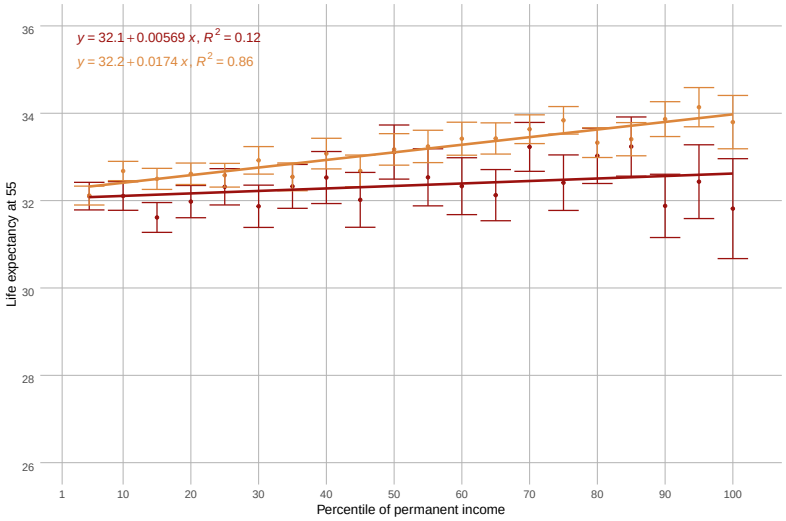
cise share of RC gap explained by each of the two factors depends on this order of decomposition. However, we consistently find in the two decomposition orders that mortality differentials account for more of the RC gap between gender than "solidarity".

Figure C.4: Life expectancy at age 55 by income group and gender – 2009, 2019

(a) Men



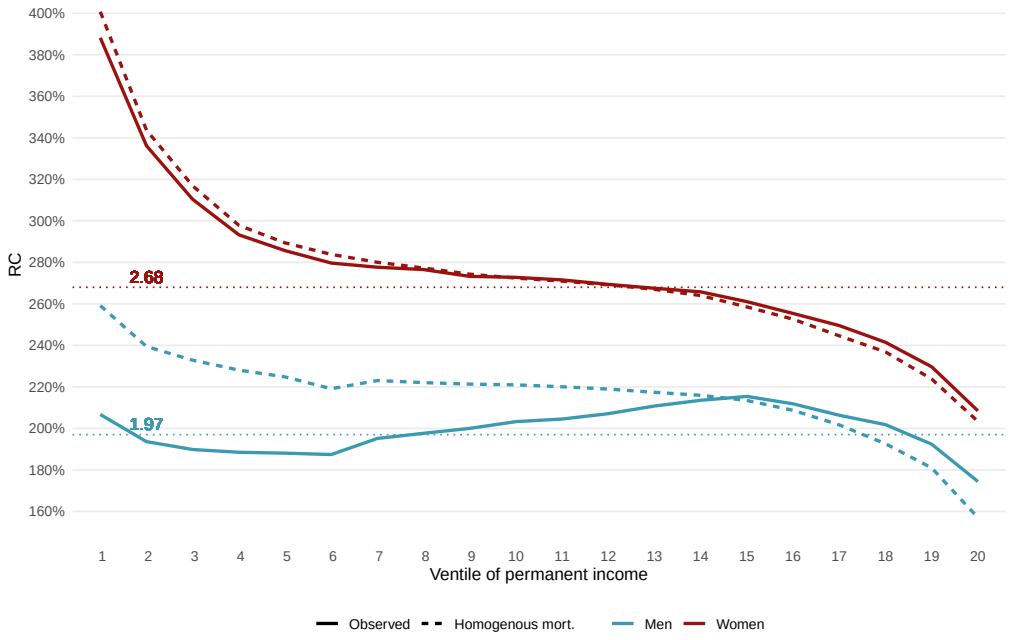
(b) Women



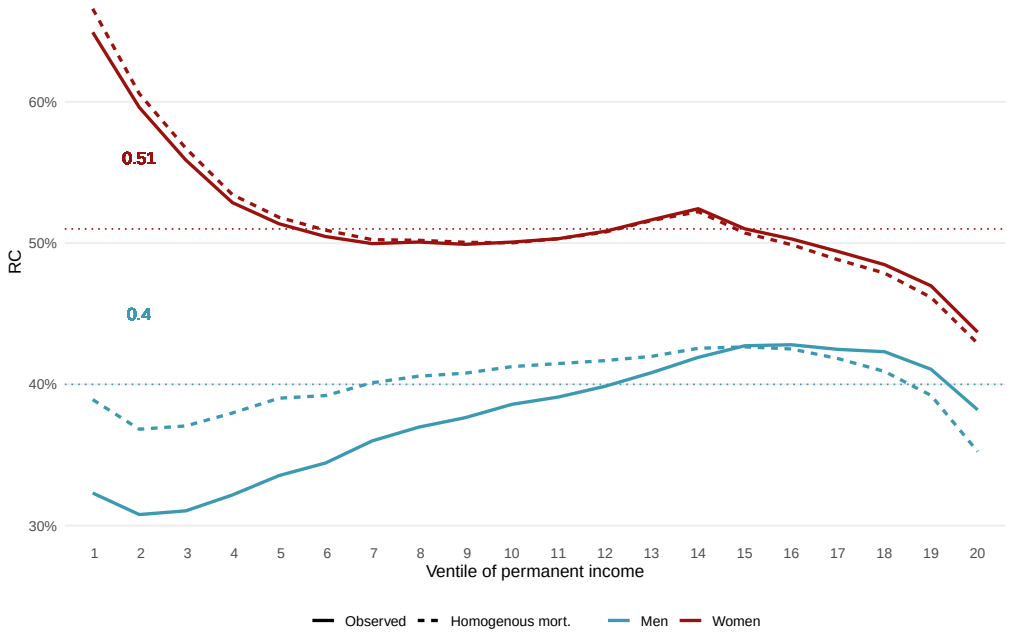
NOTE: This figure presents life expectancy at 55 by ventile of permanent income – average labour income between 40 and 54, for year 2009 and 2019 separately. Income rank is defined for the whole sample – men and women pooled. ventiles are multiplied by 5 for comparability with previous results. 95% confidence intervals follow the methodology of Camarda (2015). The level of income gradient is estimated by the regression of life expectancy on the income group.
INTERPRETATION: Between 2009 and 2019, life expectancy of men of the first ventile increased by 1.2 years. Life expectancy of men of the top ventile increased by 2 years.
SOURCE: Agirc-Arrco data, authors' calculations

Figure C.5: Decomposition of the RC gap within gender – by discount factor

(a) $\beta = 1$



(b) $\beta = 0.95$

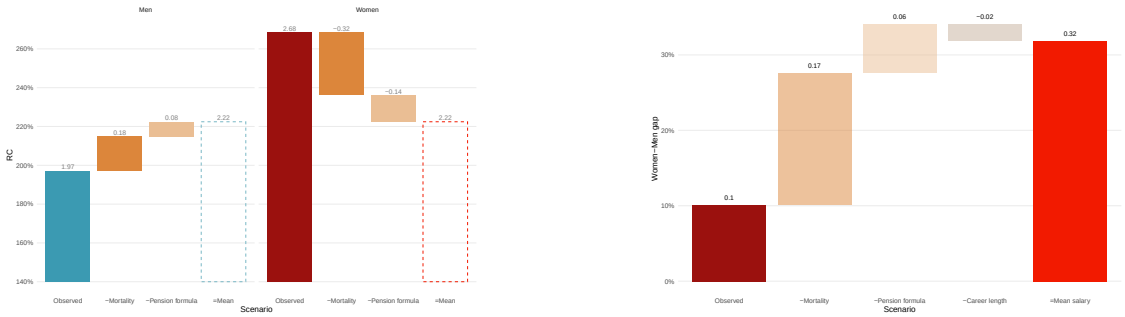


Note: This figure presents RC by ventile of permanent income – average labour income between 40 and 54 – for women (red) and men (blue), with observed (solid line) and homogeneous mortality (dashed line). Income rank is defined for the whole sample – men and women pooled. Observed RC is computed using observed contributions, pension levels, retirement age, and estimated mortality at the income ventile × gender level. Dashed line considers counterfactual returns with mean mortality rates within each gender from 55 on. Dotted line depicts average RC within gender. Discount factor equals 1 in the first panel, 0.95 in the second one.
Source: Agirc-Arrco data, authors' calculations

Figure C.6: Decomposition of the RC and SSW gap between gender – by discount factor

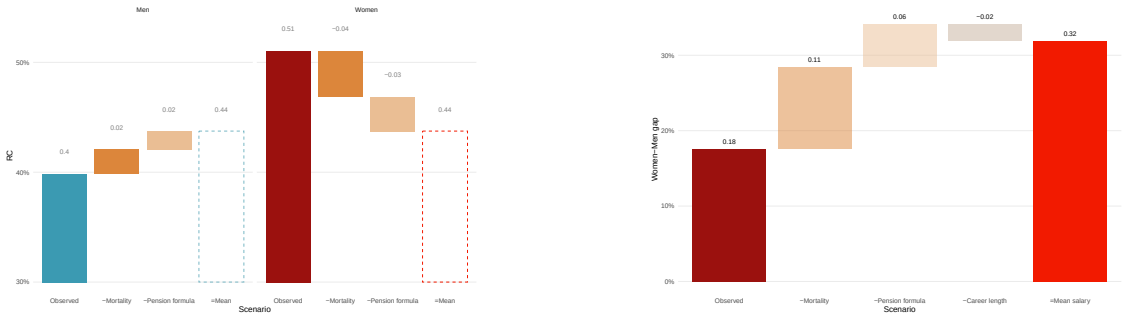
(a) RC – $\beta = 1$

(b) SSW – $\beta = 1$



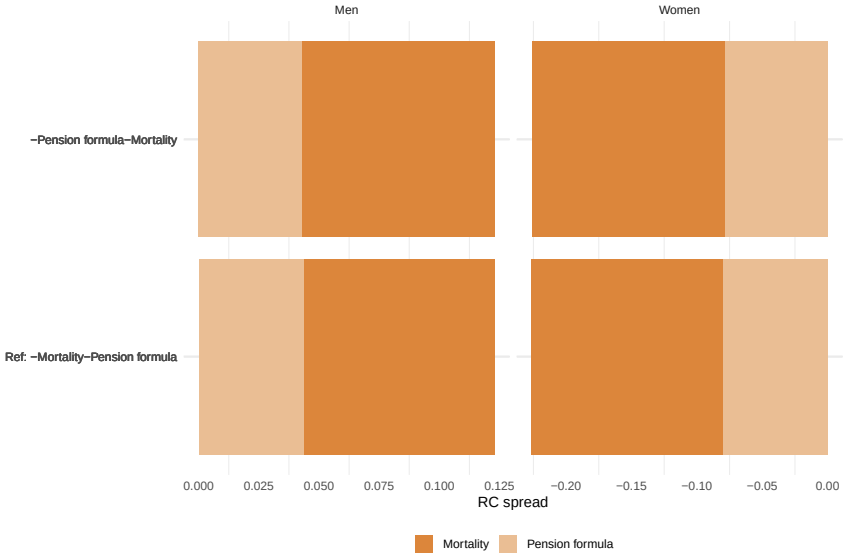
(c) RC – $\beta = 0.95$

(d) SSW – $\beta = 0.95$



Note: This figure presents average RC by gender and its decomposition – left panels – and decomposition of the lifetime pension gap between gender – right panels – using two extreme values of the discount factor – 1, 0.95.
Source: Agirc-Arrco data, authors' calculations

Figure C.7: Order sensitivity in the RC gap decomposition



Note: This figure presents the respective contributions of differentials in mortality and solidarity from the pension formula in the observed gap in RC between men and women, for two different orders of decomposition. The first line from the bottom is the benchmark scenario developed in section 4 and depicted in Figure 6. Discount factor equals 0.98.
Interpretation: In both methodologies, mortality is the main contributor to the observed gap in RC between men and women.
Source: Agirc-Arrco data, authors' calculations

Online Appendix

A Details on pension formula

In this section, we present the first two pillars of the French pension system for private workers. The two pillars are pay-as-you-go system, and compulsory.

A.1 Pension formula in the first pillar (*Régime général*)

The general formula for benefits B for wage earners of the private sector is the following:

$$B = W_{\text{ref}} \times CP \times \tau \tag{A.1}$$

The pension is proportional to a reference wage W_{ref}, which is the average of the 25 best annual earnings figures under the Social Security Ceiling. It also depends on a coefficient of proportionality (*coefficient de proratisation*, CP) accounting for the number of years contributed to the pension scheme. It is capped to one, and is computed as $CP = \max(1, D/D_{CP})$, with D the number on years of contribution in the private sector and D_{CP} a reference duration determined by law. τ corresponds to a reference replacement rate τ_{ref} of 50%, which can be either increased by a bonus in the case of continued activity beyond the full-rate age, or reduced by a penalty in the case of retirement before this age. More formally it can be decomposed as:

$$\tau = \tau_{\text{ref}} \times \Big[1 - p \times N_{pen} + b \times N_{bon}\Big] \tag{A.2}$$

Here N_{pen} is the number of quarters of penalty and N_{bon} the number of quarters of bonus. The former is calculated as the number of quarters missing to reach the normal retirement age or the full-rate duration D_{FR} (the smaller number is taken) and the latter as the number of quarters worked beyond D_{FR} and after the minimum age of eligibility (SEA). p and b respectively correspond to the rate of penalty and bonus for each quarter.

The pension accrual associated to one additional year is then the sum of three effects:

- The change in the W_{ref}, which depends on whereas the additional annual earnings is among the N best ones or not.
- The change in CP, which is equal to 0 if the reference duration D_{CP} is reached, and $1/D_{CP}$ otherwise. As this duration is the same as the full rate duration

D_{FR}, workers eligible for a full rate pension often have their CP capped to 1.

- The change in τ, which depends on the distance to the full rate age and the associated bonus and penalty rates.

A.2 Pension formula in the second pillar (*Agirc-Arrco*)

On top of the first pillar pension, former workers of the private sector get additional pension from the second pillar, often called *Agirc-Arrco* pension from the name of the organisation administrating the scheme.

The Agirc-Arrco system is point based system in which workers obtain points from their social security contributions, that are converted in pensions when the pension is claimed.

The amount of pension received by pensioners at date t is computed as:

$$P_t = \tau_{AA} \times VS_t \times \sum_{k=0}^{t} \frac{C_k}{VA_k} \tag{A.3}$$

where C_k are social security contributions to the scheme in year k. The VA_t and VS_t parameters are the parameters that respectively convert contribution into points and accumulated points into pensions at period t.

The τ_{AA} rate is corresponds to a penalty associated to the fact of claiming the pension before reaching the full rate age. The correspondence between the rate and the number of quarters missing to reach the reference duration[16] D_{CP} is piecewise linear:

$$\tau_{AA} = \min\left(1 - .01 \times (D_{CP} - D) - .0025 \times (D_{CP} - D) \times \mathbb{1}\{D_{CP} - D > 12\}, 1\right)$$

Since 2019, a temporary bonus/penalty was also included in the pension formula for workers born after 1957. The pension of a worker is multiplied by *solidarity coefficient* of .9 for 3 years (up to 67 years old) if his contribution duration is below $D_{CP} + 1$ ($D < D_{CP} + 1$). On the contrary, if a worker works 2 years beyond the target duration, he multiplies his pension by 1.1 during one year, this bonus is increased to 1.2 if he works 3 years beyond the target duration.

The scheme is governed by labour institutions and unions that annually decide of the evolution of the two parameters VS_t and VA_t. In the recent years, the two parameters more or less evolved as the same rate as prices with a few exceptions in particular years of the great recession.

[16] The target duration is the same as in the first pillar scheme.

A.3 Contribution base

The contribution base of the two pillars differ. Contribution base are defined relative to the *Plafond de la Sécurité Sociale* – Social Security Ceiling (hereafter SSC) which is an institutional threshold indexed on prices used to define contribution bases of different social benefits. In 2023, the monthly SSC was € 3,666, which is slightly more than 2 times the gross minimum wage, and a large majority of private sector workers have earnings below the SSC.

The first pillar is based on labour earnings up to one SSC, and earnings taken into account for the the reference wage W_{ref} defined above are capped up to this threshold.

The contribution base of the second pilar depends on the occupation of the worker. Agirc and Arrco were historically two separated pension scheme, one for blue and white collar workers (Arrco), and one for executives (Agirc). For non-executives (Arrco), contributions are based on labour earnings up to 3 SSC. This base is divided into two brackets: the first one that goes from 0 to 1 SSC is associated to a low contribution rate, and the second one, going from 1 to 3 SSC is associated to a higher contribution rate.[17] For executive (Agirc), the second bracket is extended up to 8 SSC from 1988 on (4SSC before). The two schemes merged in 2019, and now apply the rules formerly used for executives.

A.4 Non-contributory components

On top of the contributory components described above, the two pillars have additional solidarity components that provide pension rights to the workers despite the absence of contributions. We describe the main components below.

Implicit non-contributory components Given that it is based on points, the pension formula makes a clear link between contributions and pension. This is not the case of the first pillar based on annuity. In particular the way insurance duration is computed generates additional pension rights for people with low earnings. The duration D is actually computed based on the total earnings of the workers during the year, and not on the actual working duration. To get the number of validated quarters in year t, one divides the total earnings by gross earnings associated to 150

[17] In 2018, the low contribution rate was 6.2% of gross earnings and the high contribution rate was 16.20%. For the first bracket, one has to add the contribution rate associated to the first pillar: 15.45% of gross wages (6.9 of employee contributions and 8.55 of employer contributions).

hours worked at the minimum wage ($\underline{\mathrm{W}}$). Duration in year t is thus D_t:

$$D_t = \min\left(\left\lfloor \frac{W_t}{\underline{\mathrm{W}}} \right\rfloor, 4\right)$$

where $\lfloor . \rfloor$ is the floor operator. 150 hours of work correspond to about one month of full time work, so anyone who works 4 months full time at the minimum wage will validate one year of insurance duration.

Pension rights associated to children First, maternity leave and education of children provide additional pension rights.

- In the first pillar, the birth of each child provide 4 quarters of contribution for the mother. 4 additional quarters can be shared between the parents for the education of children. By default, these 4 quarters are given to the mother, but parents can ask to change this allocation. In addition to that, the pension of parents of 3 children or more is increased by 10%.

  Moreover, parents who take a parental leave (either stop or reduce their working activity to educate their children) benefit from additional pension rights that are equivalent to working time at the minimum wage. It thus increase pensions through both the reference wage (W_{ref}), and the number of years of contributions (D).

- In the second pillar, if the number of kids is observed at the moment of the pension claim, the bonus applied to kids depends on the period of accumulation:

  - For non executives (Arrco), points accumulated before 1999 are inflated by 10% if the worker has 3 kids, and then by 5% per additional child.[18] Between 1999 and 2011, the bonus was reduced to 5% for parents of 3 children or more, and there was no additional increase with the number of children.
  - For executives (Agirc) points accumulated before 2012 are inflated by 8% for parents of 3 children, and then by 4% per additional child.
  - The bonus was standardised for executives and non executives since 2012: parents of 3 children or more get a 10% bonus on their pension. The total bonus is capped to a threshold.

[18] A parent of 3 kids had a 10% bonus, a parent of 4 kids a 15% bonus, etc.

Unemployment, sickness and disability periods

- In the first pillar, unemployment, sickness and disability periods are taken into account in total insurance duration D, but do not contribute to the reference wage W_{ref}. Information about the situation of the worker is usually transfered by the administration in charge to the pension scheme.

- In the second pillar, points are attributed to workers based on their previous earnings:

  - If the worker gets a disability benefit, sickness leave, maternity leave benefit, the workers gets additional points to his pension. Average points per are computed based on the year preceding the leave, and multiply by the number of days of leave to obtain the number of point given to the worker.

  - Unemployed workers gets additional points that are proportional to the unemployment benefit they perceive.

Indexation rules Finally, indexation rules may also create distortions between contributions and pensions. *Per se*, this does not constitute a non-contributory component. However, it is worth taking that into account given that it may affect return rates of workers.

In the first pillar, indexation rules play an important role in the computation of pension through the reference wage. Reference wage is composed of the best 25 years, and past earnings are indexed on prices. The fact that these are indexed on prices rather on the total payroll[19] has important redistributive implications: workers with more increasing earnings benefit more from the system than others.

A.5 Early start exceptions

The 2003 reform of the pension system progressively increased the target duration from 37.5 to 40 years of contribution. The same reform allowed workers who started to work before 17 and that reached the target duration to claim their pension before the early retirement age, that was equal to 60 years old at that time. The 2010 reform that progressively increased the early retirement age from 60 to 62 years old extended these exceptions: for the most recent generations, workers can claim

[19]The total payroll evolution is the natural return rate of any pay-as-you-go pension system. See Samuelson (1958).

their pension from 58 if they started to work before 16 years old, or from 60 if they started to work before 20 years old.

One has to keep in mind that these exceptions are conditional on having reached the target duration (43 years for the most recent cohorts), which basically means that they have been working continuously since their career start.

Workers claiming their pensions through these exception benefit from the full rate in both the first pillar ($\tau = .5$), and the second pillar ($\tau_{AA} = 1$). No temporary bonus/malus is applied to their second pillar pension.

B Details on data sources

B.1 Description of the different data sources

***Agirc-Arrco* records** We can categorize *Agirc-Arrco* administrative records into three main categories: first, the *Base individus* presents an index of all *Agirc-Arrco* affiliates, be it for direct or indirect rights. This database is organised with one observation per individual and presents many summary variables derived from other administrative records. We use it for some general information – gender, date of birth and death, location of birth, direct or indirect rights.

Second is the career *RNGD* database, that breaks down earnings and the acquisition of direct rights in the career spent in the *Agirc-Arrco* regimes – one observation per individual-year-pension regime (*Agirc*, *Arrco*, *Agirc-Arrco* after 2019). The main variables of interest for our analysis are points per period, pension regime, contribution bracket and type of points – work, disease etc – as well as related earnings. We also recover time spent in the regime from data on entry and exit date in the regime, for each period and type of state – work, unemployment etc. In principle, all years of the *Agirc-Arrco* career of individuals with direct rights should be present. In practice, this data suffers from a lack of completion for old years or cohorts – see next subsection.

Third, *Agirc-Arrco* records provide information on pensioners in different databases. We first use *Bases allocataires* – pensioners database, one per year between 2009 and 2022, that contains information on total points validated at claiming, by regime (*Agirc*, *Arrco*, unified *Agirc-Arrco*) and for three large categories of points: free points for children born and raised, free points for seniority, and other points. For each year, *Bases allocataires* contain individuals who are pensioners of the regime during the year, and possibly those who died the year before. We also recover for these the penalty rate used in the computation of *Agirc-Arrco* pensions – see Online Appendix A. From *RCIV* records, we recover total periods validated by pensioners of the regime up to 2022. These periods are differentiated between those validated in the *Agirc-Arrco* regime and other regimes, broken down by type of validated quarter and year. From the *RCIV* database, we also recover claiming dates for the first- and second-pillar. Finally, the *EIRR* database allows us to recover pension amounts received by pensioners for all regimes – in particular *CNAV* and *Agirc-Arrco* pensions. One database is built by quarter and we have access to the Q1-2014 to Q4-2022 databases. Each quarter's "*EIRR*" contains pensioners alive during the quarter – possibly some who died in the previous quarter. From this last data source, we use pension amounts for *CNAV* and *Agirc-Arrco* and the date

of last revalorisation of pension amounts to compute pension level at claiming.

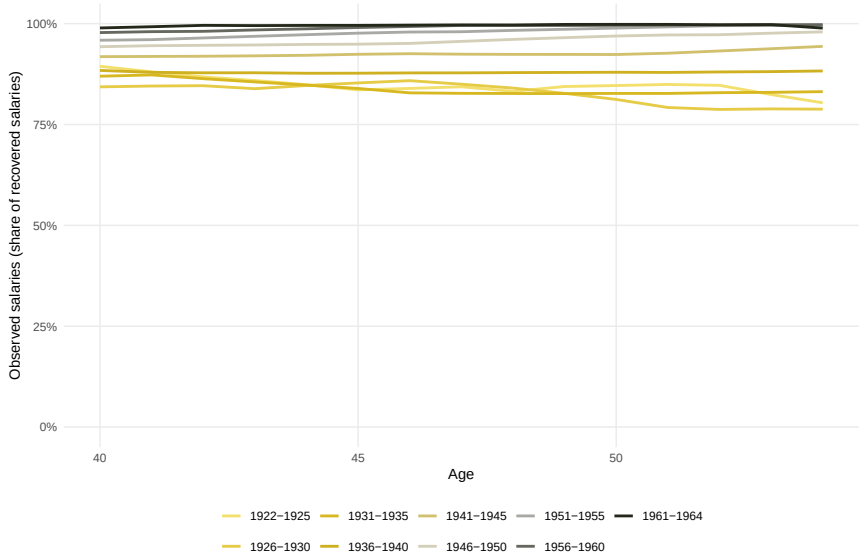
Other data sources Apart from administrative records, we use various sources to recover main parameters of the pension system, used in the computation of permanent income and IRR. Some come from *Agirc-Arrco* archive documents (value of the Social Security Threshold, inflation index – IPC, mean wage index – SMPT, purchasing and service values of points, adjustment rates). From the *IPP tax and benefit tables* (2023), we recover historical values of the minimum wage and of minimum and maximum contribution rates for the first- and second pillars. We also leverage average contribution rates, as computed in a working paper of the French Retirement Guidance Council (COR, 2009) for the 1963-2009 period. Before 1963 and after 2009, we choose the maximum value between 1963 or 2009 mean rates as computed by the COR, and contribution rates indicated in the *IPP tax and benefit tables* (2023). Finally, to compute net wage in the sensitivity analysis, we mobilise series of purchasing power of the average net wage, from the National Statistics Institute – *Insee* (Insee, 2023a,b).

B.2 Quality of the information on careers

The *Agirc-Arrco* data on careers poses several challenges for our analysis. The first one relates to completion of career data. In theory, all periods spent in the *Agirc-Arrco* regime during the career are informed in the *RNGD* database, be it for accumulated points, earnings or duration spent in the regime, for all present and past affiliates. In practice, this is not the case, because of regularisation of career information at claiming and lack of completion for old periods. It poses a challenge for computing permanent labour income out of incomplete career data. We use earnings data when available. However, Figure B.1 shows that for older cohorts, earnings data is more incomplete than points data: ratio of total earnings over recovered earnings from point is lower than 100%, increasing with birth year and relatively stable across ages of the career within a cohort. When points are available and earnings are not, we reconstruct salaries from points, using the price of points and estimates of individual contribution rates. A related difficulty is that, by definition, we only observe earnings for periods spent within one of the *Agirc-Arrco*. labour income earned in other sectors – *e.g.* the public sector – are therefore not observed. For these two reasons, in our permanent income computation, we normalise the sum of labour income by the sum of days spent in the sector, to prevent bias from individuals with high income who spent only few years in the private sector. We also restrict the analysis to individuals who spent at least 3

years in the sector between 40 and 54, for whom we assume observed income is more or less representative of earnings on the whole career. In section 4.1, we found that life expectancy among men of the first quantiles is oddly high, suggesting they might benefit from other sources of income. Restricting to individuals of the sample from cohort 1954 who retired and for which we have data on pensions, we do find evidence of this: individuals of the first five ventiles are relatively low in the distribution of *Agirc-Arrco* pension, but relatively high in the distribution of total pension, suggesting that a principal share of the career was spent in other sectors. This result is found in Figure B.2. Similarly, in Figure B.3 among cohorts 1922 to 1964 for which pension is observed, we find that on average, less that 80% of total pension is made of private sector regimes among men of the bottom 7 percentiles.

Figure B.1: Ratio of observed earnings over all recovered earnings



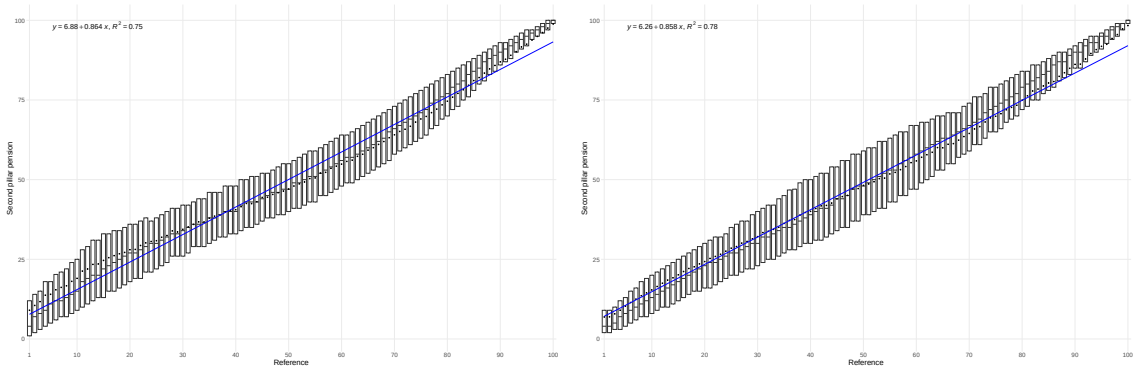
NOTE: This figure presents total salaries from earnings database over total recovered earnings by age and cohort group. Total recovered earnings equal observed earnings when available and earnings recovered from points when not. These are obtained by multiplying points by purchase price and dividing by individual contribution rates, following methodology detailed in B.1.
INTERPRETATION: For most recent cohorts, total observed earnings are almost equivalent to total recovered earnings. Among older cohorts, observed earnings are more incomplete.
SOURCE: Agirc-Arrco data

Another challenge posed by earnings and points data is that recovered earnings are top coded. Indeed, there are two contribution thresholds above which additional earnings give no additional points: 3 times the Social Security Ceiling – SSC – for non-executives, 4 times for executives before 1988, 8 times after. For consistency across cohorts and years, we compute permanent income using earnings top-coded at 4SSC. However, to compute lifetime contributions for the 1954 cohort, we retain all contributions up to 8SSC. In Figure B.4, we display, for cohort 1954, the distribution of observed earnings (brown, top-coded at 8SSC after 1988) and the same

Figure B.2: Quantile-Quantile plot: Permanent labour income and pensions

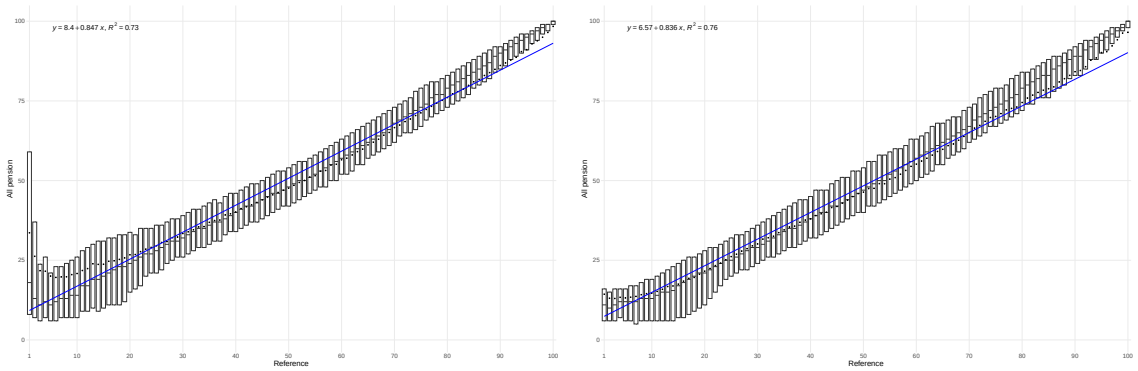
(a) Agirc-Arrco pension, Men

(b) Agirc-Arrco pension, Women



(c) Total pension, Men

(d) Total pension, Women



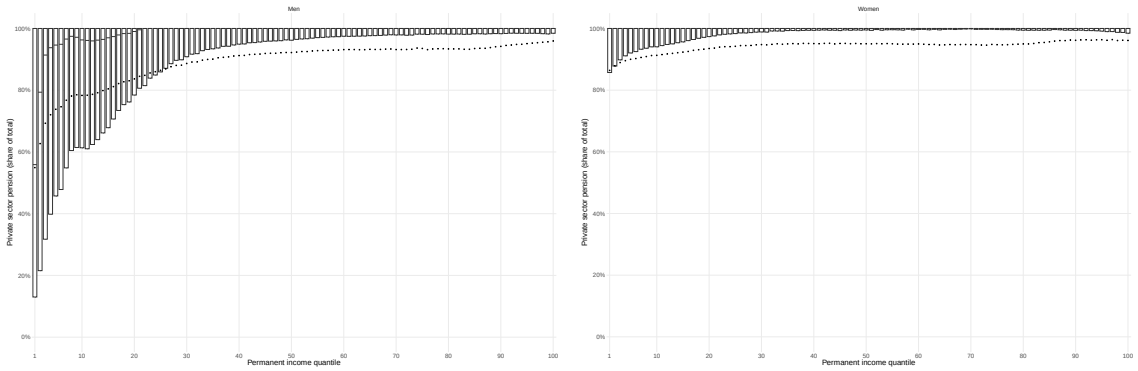
NOTE: This figure presents the observed distribution statistics of pension quantiles, by percentile of permanent labour income. For each quantile of permanent income, horizontal tick, diamond and box respectively represent median, average and interquartile range of pension percentile. Sample is made of individuals of section 4.2 born in 1954 for which pension level is observed. Pension is *Agirc-Arrco* pension in the first panel and total pension in the second one.
INTERPRETATION: Men of the first permanent income percentile are on average at the 26^{th} percentile of total pensions.
SOURCE: Agirc-Arrco data

distribution, excluding the top contribution bracket (from 4 to 8SSC). It appears that bunching at 4SSC concerns less than 1% of observations. Top-coding is therefore unlikely to have significant impact on permanent income ranking, even at the very top of the distribution.

Finally, let us note that even though data on points is less incomplete than earnings data for older cohorts, it is still rather incomplete for this population, as apparent in Figure B.5: total points in the career database do not come close to total points claimed at retirement for cohorts that retired before 2009 – even less before 1999. In the computation of permanent income, our main strategy is to normalise earnings by number of days worked to circumvent this issue. Number of days worked have the same data source as points in the *Agirc-Arrco* administrative records, and their respective levels of completion are the same. We also notice that for older cohorts, data is more incomplete for older years. This is apparent in

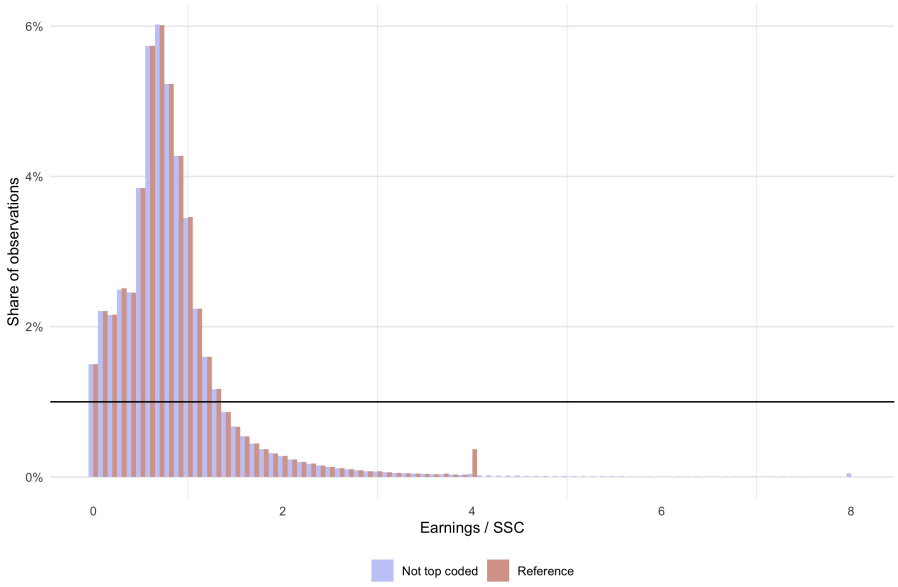
Figure B.3: Share of pension from the private sector, by permanent income percentile

(a) Men (b) Women



NOTE: This figure presents the observed distribution statistics of the share of private sector pension over total pension, by percentile of permanent labour income. For each quantile of permanent income, horizontal tick, diamond and box respectively represent median, average and interquartile range of pension percentile. Sample is made of individuals of section 4.1 for which pension level is observed, before truncation of the first 7%. Private sector pension is defined as *Agirc-Arrco* and *Cnav* pension.
INTERPRETATION: On average, private sector pensions represent 60% of total pension among men of the first permanent income percentile.
SOURCE: Agirc-Arrco data

Figure B.4: Distribution of earnings for the 1954 cohort

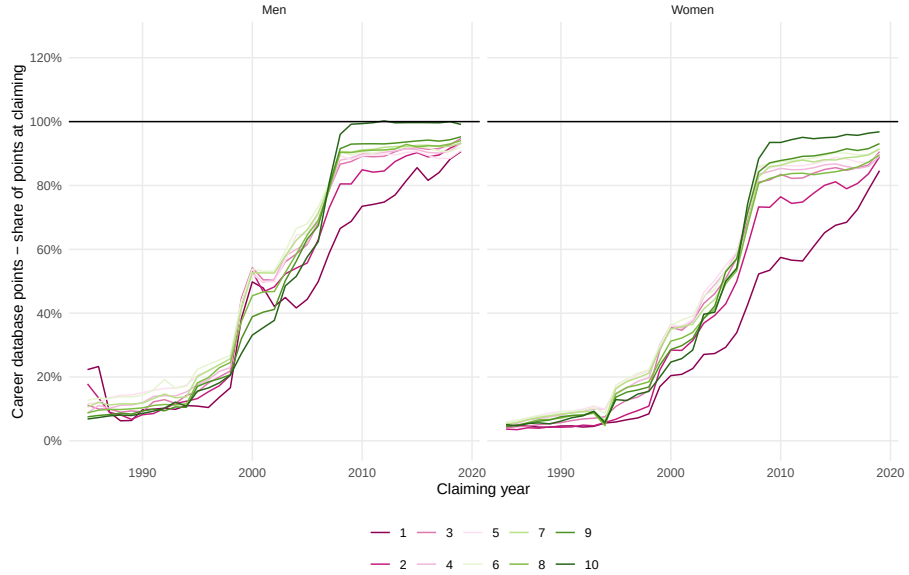


NOTE: This figure presents the distribution of observed earnings, normalised by the value of the Social Security Ceiling – SSC. Observed earnings (top-coded at 8 SSC) are presented in brown. Observed earnings excluding the top contribution bracket are presented in purple. The sample is made of one observation by individual-year with positive accumulated points and positive earnings in the career records, for sample born in 1954 used in section 4. Binwidth equals 0.1 SSC.
INTERPRETATION: Less than 1% of recovered annual wages are bunched around 4SSC or 8SSC.
SOURCE: Agirc-Arrco data

Figure B.6: for older cohorts, earnings of the end of the career are overrepresented in career history, since data is more incomplete at the beginning of the career. This is one of the reasons why we restrict our analysis to a narrow age window

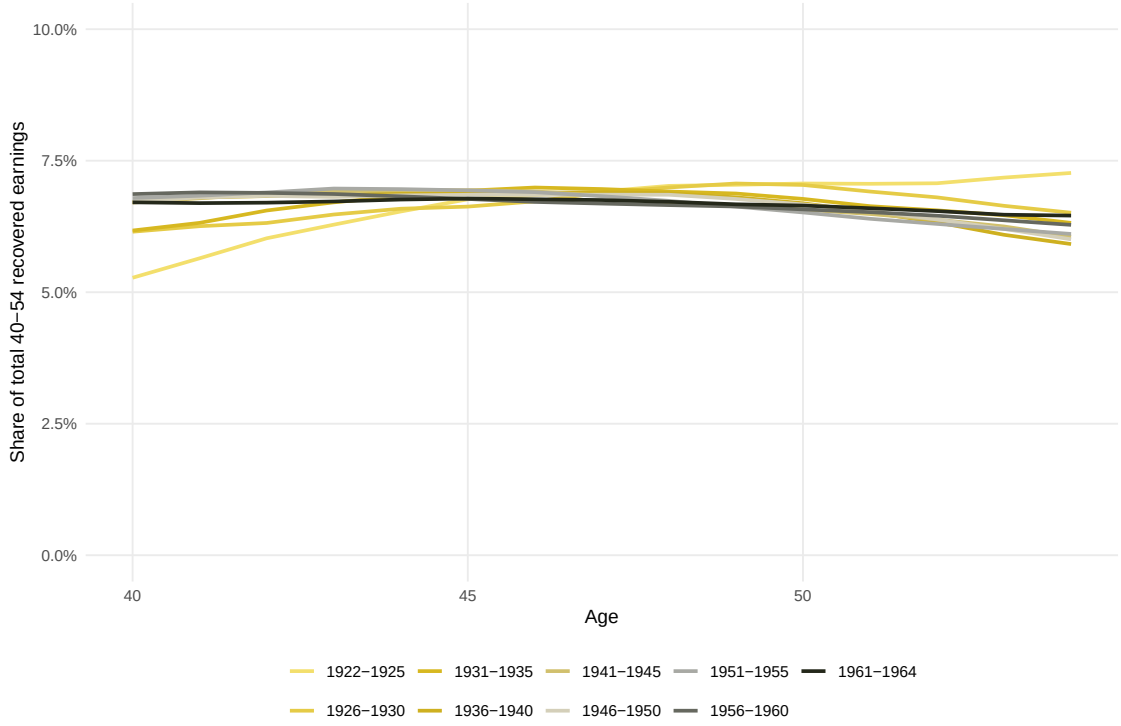
between 40 and 54 to compute permanent income, test even narrower windows, in Appendix C.1, with low impact on our results.

Figure B.5: Data quality on accumulated points: comparison of the career and pension records



NOTE: This figure presents the ratio of total accumulated points in the career records on the same measure in the pension records by claiming year, gender, and permanent income decile. Income rank is defined for the whole sample – men and women pooled. Population is made of individuals of the sample used in section 4.1, who claimed their pension at some point between 1985 and 2019 and had a minimum of 100 points at claiming. All points are taken into account, exception made of "free" points at claiming for seniority of raised children.
INTERPRETATION: On average, for men of the 2^{nd} decile who retired in 2009, 20% of total points at claiming are missing in career records.
SOURCE: Agirc-Arrco data

Figure B.6: Split of 40-54 recovered salaries, by age and cohort



Note: This figure presents the split of total earnings between 40 and 54, by age and cohort group. Salaries are normalised by an index of mean salary – *SMPT*. Sample is the one used in section 4.1.
Interpretation: For most recent cohorts (after 1950), the distribution of earnings and periods of work across the career is relatively flat between 40 and 54 (around 1/15 of total for each year, during 15 years). For old cohorts, data on young ages are under-represented in the career records.
Source: Agirc-Arrco data