IFS

Institute for Fiscal Studies

Aurel Mélard
Simon Rabaté
Maxime Tô

Working paper

24/44

Hidden redistribution in lifetime earnings: the role of differential mortality

Economic and Social Research Council

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

Aurel Mélard† Simon Rabaté‡ Maxime Tô §

October 2024

Abstract

Differences in life expectancy between gender and income groups are large and generate significant implicit redistribution in lifetime earnings through pension systems. We use administrative data on the universe of private sector wage earners in France to quantify these effects. We establish two main results. In terms of between gender redistribution, we find that differential mortality – life expectancy at 55 is 5.7 years higher among women than among men – reduces lifetime income inequality between men and women: absent this differential in life expectancy, the pension gap between men and women would be 72% higher, in a lifetime perspective. Second, within gender, high income earners benefit from hidden lifetime redistribution due to higher life expectancy. We find a life expectancy gradient of 7.2 years between the extremes of the distribution among men, 1.8 years among women. Among men, this hidden redistribution more than offsets the overall progressivity of the pension system.

*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, Johannes Spinnewijn, Julie Tréguier and Mickaël Zemmour, as well as participants to the PENSINEQ meetings for their valuable comments and suggestions. Remaining errors are our own. This research was funded by Agirc-Arrco, and by the Joint Programming Initiative (JPI) "More years, better lives" (PENSINEQ Project, ANR-21-MYBL-0002).

†CREST, École Polytechnique de Paris

‡Ined and Centraal planbureau

§Institut des Politiques Publiques, CREST and Institute for Fiscal Studies

1 Introduction

Over the last decade, the study of income inequality has made enormous progress, shedding light on the magnitude, evolution, and determinants of inequality between income groups (Piketty et al., 2018). We also have a better understanding of how these differences are attenuated by tax and transfer systems (Bozio et al., 2024, Bruil et al., 2022). Simultaneously, research on lifetime earnings inequality has advanced significantly (Kopczuk et al., 2010, Guvenen et al., 2022, 2021), revealing the importance of considering long-term economic trajectories. However, most studies consider instantaneous redistribution, either at one 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: richer 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. On one hand, high-income individuals receive their relatively high pension benefits for a longer period, potentially exacerbating lifetime income inequality. On the other hand, women's pensions, while typically lower due to various 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 a longer period. This may reduce the gender gap on a lifetime basis. Given that pensions constitute the largest public expenditure in many countries,[1] 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 gender and within gender between income groups.

This paper aims to address this crucial gap in the literature by quantifying these effects and providing a more comprehensive understanding of lifetime inequality and redistribution, taking into account the particularities of gender differences in lifetime earnings and life-expectancy.

Our study quantifies the impact of differential mortality on redistribution through pension systems. We examine this effect both between genders and across income groups within each gender, offering a comprehensive view of how longevity differences shape lifetime redistribution. Focusing on wage earners in the French private

[1] Dherbécourt et al. (2023) show that in Europe, pensions account for close to 50% of social protection spending.

sector, we leverage administrative records covering the entire population of affiliates to produce two key sets of findings.

First, we establish the relationship between income and life expectancy. Using permanent labor income, 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 find a clear income gradient in life expectancy, nearly linear in income percentile and concave in income level. This gradient is more pronounced for men (seven years, with a population average of 27) than for women (two years, with a population average of 33).

Second, we analyse how differential mortality affects the overall redistribution generated by the pension system. We compare the return on contributions (RC) across income group and genders, with and without differential mortality. The RC compares benefits received from the pension system to contributions paid, and has a direct interpretation in terms of redistribution : individuals or groups receiving above (resp. below) average RC are net beneficiaries from (resp. contributors to) the system. We show that, absent differential mortality, RC decreases with income, reflecting the system's progressivity. However, this pattern masks substantial between and within gender redistribution that we analyze in turn. Between genders, we find large redistribution from men to women, with women's average RC 37pp (35%) higher than men's, mostly due to women's higher longevity. Absent the differences in longevity, the gender gap in lifetime pensions would be 13pp (72%) higher. Within genders, differential mortality almost entirely reverses the RC progressivity for men. High income men receive significant implicit transfers : those in the top vintile receive a 10% pension subsidy due to longer life expectancy, corresponding to an implicit tax of up to 18% on low income men's contributions. For women, consistent with their smaller mortality gradient, RC remain relatively stable across mortality scenarios.

This paper contributes to four 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.[2] 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 (Dahl et al.,

[2]Chetty et al. (2016), has been replicated in different countries: Kinge et al. (2019) for Norway, Dahl et al. (2021) for Denmark and Blanpain (2019) for France.

2021). We also focus on individual labor earnings rather than household income, a choice driven by data constraints but also relevant given that labor earnings form the tax base for pension systems. These methodological differences yield a smaller income-mortality gradient compared to Chetty et al. (2016), likely due to our different income concept.

Second, we contribute to literature on the distributive effects of pension systems, particularly their interaction with differential mortality. While static analyses of pension progressivity are common (Whitehouse, 2006), accounting for differential mortality requires comprehensive data on income and mortality trajectories. Previous studies using microsimulation models (Liebman, 2002, Mazzaferro et al., 2012, Dubois & Marino, 2015) found moderate effects of differential mortality on redistribution. However, these studies were limited by their reliance on survey data and education-based mortality estimates. Our work, like that of Haan et al. (2020), overcomes these limitations by using administrative data. We extend their approach by including women in our analysis and quantifying the relative magnitude of the mortality effect compared to other sources of lifetime redistribution.

Third, we contribute to research on gender gaps in pensions. This literature typically focuses on cross-sectional comparisons of pension level (Betti et al., 2015, Samek Lodovici et al., 2016), examining factors such as gender differences in work hours (Bertrand, 2020), hourly wages (Blau & Kahn, 2017) which translate into lower pension rights. There are also differences between men and women in terms of saving (Neelakantan & Chang, 2010) or retirement decisions (Tréguier, 2021), which can be related to differences in risk aversion or social preferences (Bertrand, 2011). Our study adds to this 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 factors.

Finally, we contribute to a broader literature on the effects of the tax and transfer system on income inequality. Recent work following Piketty et al. (2018) has provided comprehensive analyses of inequality and redistribution (Alvaredo et al., 2020, Bozio et al., 2024, André et al., 2023). However, these approaches often neglect lifetime redistribution, that can be high in social insurance schemes such as unemployment, disability, and most importantly pension system. We show that the pension system system in general and differential mortality in particular generates significant monetary transfers between low and high income, which highlights the importance of considering lifetime effects in overall assessments of redistribution.

The rest of this paper proceeds as follows. Section 2 describes the institutional context and the data sources used in the analyses. We then present the results on the relation between income and mortality in section 3 and the analyses on the redistribution generated by the pension system in section 4. Section 5 concludes.

2 Institutional context, data and methodology

2.1 Institutional context

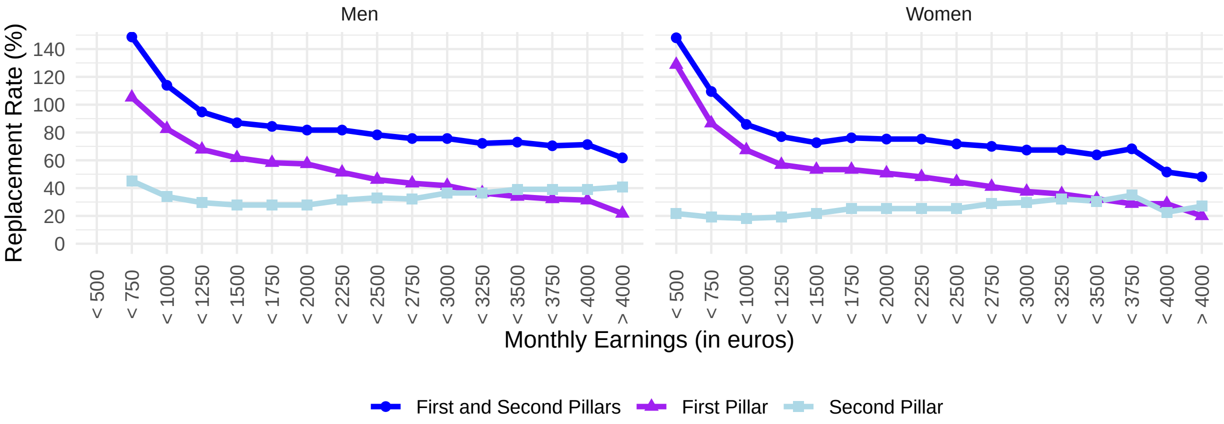
The public-pension system in France is large, with benefits amounting to roughly 14% of GDP. In this paper, we focus on wage earners of the private sector. For this group, the pension is divided into two components that are provided by two different pension schemes : a first pillar (*Régime général*) and a second pillar (*Agirc-Arrco*). A high share of the French working population is affiliated to these regimes at least once in their career : 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 base, their pension formulas – annuities for the first pillar, point system for the second, and their governance – state for the first pillar, labor organisations for the second.[3]

The two pillars are contributory, in the sense that pension benefits are increased with pension contributions. However, the pension formula and the way solidarity components (accounting for births, disability, unemployment, etc.) are accounted for in the system may lead to important differences in returns on contributions. Figure 1 taken from Aubert et al. (2012) shows that the replacement rate (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 that incorporates more solidarity components in its formula. In contrast, the second pillar replacement rates do not vary much according to labour earnings reflecting a stronger contributory feature of the second pillar. Aubert & Bachelet (2012) show that redistributive aspect of the first pillar is mainly driven by the solidarity components associated to parentality or compensating career interuptions. The basic pension formula is actually anti-redistributive : when neutralizing all solidarity components, the D9/D1 ratio of pension is larger than the D9/D1 ratio of cumulated earnings.

[3] Appendix C presents in details the main features of the pension formulas for the two schemes.

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

Administrative records from Agirc-Arrco The primary data sources for this study come from the administrative records of the*Agirc* and *Arrco* pension schemes. These provides detailed information about the universe of the affiliates to the scheme. Demographic data includes individuals' dates of birth and death, gender, and country of origin (France or abroad).

This data also offers detailed information on individuals' earnings trajectories in the private sector, from 1962[4] 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).

We use information on points to compute average earnings during the career (see appendix C for the link between earnings and contributed points). Our analysis also requires to observe the level of contributions for both the first and second pillars. Contributions to the second pillar are directly recovered from acquired points and their purchase value in *Agirc-Arrco* records, while contributions to the first pillar can be inferred from earnings using historical contribution rates.

In addition to career information, the *Agirc-Arrco* records provide exhaustive information on the second pillar pensions, with many details on the different com-

[4]The Agirc scheme was created in 1947, and the Arrco scheme in 1962. We thus have comprehensive data since 1962.

ponents of these rights – direct 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 affiliates who retired from 2014 onwards. We use this information in section 4, when studying the overall redistribution in the French pension system for the private sector (first and second pillars combined).

This rich data enables us to analyse redistribution from a lifetime perspective, based on observed earnings during the career and pension during retirement. However, it is important to note that data quality improved over time, especially regarding career data : career trajectories might be incomplete for less recent years, particularly for individuals who died before 2009, retired before 1999, or more generally for older cohorts (see appendix B). We address this issue in our choice of sample and income concept, as discussed in the next section.

2.3 Definition of permanent income and sample selection

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 labor 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 labor income to total amounts of pensions perceived. On the other hand, using permanent income rather than lagged income as a determinant of mortality potentially limits the issues raised by income mobility during the life cycle (see Dahl et al., 2021).

Sample selection Ideally, we would like to rank all affiliates of the system using total labor earnings over the life-cycle. We detail below how we deviate from this ideal setting in terms of sample selection and definition of permanent income.

The initial sample consists of all *Agirc-Arrco* affiliates alive in 2009. We also restrict our sample to people who are observed after 55 years old, as we measure life expectancy at that age. Affiliates are contributors or pensioners who acquired direct rights in the scheme (ie. employees or former employees from the private sector) and thus appear in the administrative records.

Among all affiliates we exclude individuals whose contributions do not appear in the records or only before 1962 (creation of *Arrco*, sub-scheme for non-executives). Before this date, only executives contributed in the scheme. We thus exclude individuals who only contributed before 1962, and some pensioners whose contributions are not reported in the earnings records, as data quality is lower for older cohorts.

We then focus on France-born individuals, due to measurement issues for both earnings and mortality among individuals born abroad : if a large share of the career is spent abroad, our measure of earnings might not reflect actual lifetime labor income. Mortality is also poorly measured for this population when these affiliates do not live in France at time of death.

Our sample selection is also driven by our definition of permanent income. We choose to define permanent income based on average earnings observed between 40 and 54. First, since we analyse mortality starting at age 55, the upper threshold of the age window needs to be below this age to avoid reverse causality between income and mortality. Focusing on relatively old ages allows us to use more recent and better quality data. Thus, we select individuals who are affiliated to the *Agirc-Arrco* scheme at least once between 40 and 54. This age window determines the cohorts that we study : those who were 40 after 1962 and 55 no later than 2019 (last year of our study). We therefore focus on the 1922 to 1964 cohorts.

Finally, we want to capture a measure of earnings that is as representative of life-cycle income as possible. We therefore exclude individuals who contributed to the private sector scheme for less than 3 years between 40 and 54, as their average private sector income is likely to be poorly correlated with total life-cycle income.

Table 1 summarizes the impact of restrictions on sample size. Starting from 14.3 million men and 12.3 million women, we end up with 5.2 million men and 4.6 million women in our final sample. These restrictions have a significant impact on the sample size. However, it should be noted that most the important restrictions tend to exclude individuals who are not working much in the French private sector : they may be workers from other sectors (like the public sector) who occasionally contributed to private sector, or people out of the labour force between ages 40 and 54. We discuss in Appendix D the sensitivity of our results to alternative sample definitions. They appear robust to these robustness checks.

For the selected sample, permanent income is defined as average labor income between ages 40 to 54. Earnings are summed over these years and normalized by the number of days worked in the *Agirc-Arrco* scheme during this period. This normalization limits the bias linked to individuals with high revenues but a short share of their career in the private sector. Earnings are also normalized by an

Table 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
Reported contribution ≥1962	M	12.2	85	
	F	10.2	82	45.4
Born in France	M	9.2	75	
	F	8.4	83	47.8
Worked once between 40 and 54	M	6.3	69	
	F	5.9	70	48.1
>3 years in the AA scheme	M	5.2	83	
	F	4.6	78	46.5

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

index of mean wage per capita : earnings between 40 and 54 are made comparable across years and cohorts to define income ranks across all cohorts. Details on computation of permanent income can be found in Appendix C.3. In what follows, unless specifically mentioned, quantiles of permanent income are defined pooling men and women (same quantiles for both). Finally, as mentioned in Attanasio & Hoynes (2000), the definition of income quantiles among our sample is exposed to the risk of survival bias among the cohorts that we only observe at ages older than 55 : for example, cohort 1944 is 65 in 2009, the first year of mortality observation. If the top percentile of the income distribution lives longer on average than the rest of the population, they represent a higher share than 1% of the remaining population at 65. In our main estimates, quantiles of permanent income are therefore defined for the whole sample based on the 1964 cohort's permanent income distribution, the "youngest" of the sample that we observe at 55. We test an alternative definition of quantiles by cohort in Appendix D.

3 Differential life expectancy

3.1 Measuring life expectancy

Mortality is defined using cross-sectional data on observed deaths and survival of cohorts 1922-1964, averaged over the 2009-2019 period. Like in most of the aforementioned literature on mortality inequality, we use a parametric model of mortality to smooth observed mortality rates. More precisely, we estimate a logistic model using observed mortality rates Q_a between ages 55 and 90:

$$\ln\left(\frac{Q_a}{1-Q_a}\right) = \alpha + \beta a + \varepsilon_a$$

Smoothed mortality rates at age a, $\tilde{Q}_a$, are obtained as the predicted values of the model.

Above 90 and up to 120, mortality rates are those of the overall French population, as measured by the national institute for statistics on the 2009-2019 period. Survival at age 54 is equal to 1 by definition and from age 55 to 120:

$$\forall\ a \geq 55\ :\ S_a = \prod_{i=55}^{a}(1-\tilde{Q}_i)$$

Life expectancy at 55 is then equal to the sum of surviving probabilities at all subsequent ages ($S_{54} = 1$ and $S_{121} \approx 0$):

$$\begin{aligned} E_{55} &= \frac{55(S_{55}-S_{56}) + 56(S_{56}-S_{57}) + \ldots + 120(S_{120}-S_{121})}{S_{54}} \\ &= \sum_{a=55}^{120} S_a \end{aligned}$$

Survival rates and life expectancy are two of our main variables of interest in what follows. In the subsequent analysis, we compute them conditional on gender and income group.[5]

3.2 Life expectancy gradients

Figure 2a and 2b respectively show the life expectancy at 55 for men and women by permanent income percentiles.[6]

Male life expectancy gradient For men, we find a 7.2-year gradient in life expectancy across permanent labor income. Life expectancy generally increases with permanent income, except between the 1^{st} and 7^{th} percentiles. These income groups incorporate individuals whose labor income in the private sector represents a small share of total labor income, and for whom other sources of income are likely linked to higher longevity. In Appendix B, we show that those individuals indeed

[5]More details on the methodology and sensitivity analysis are provided in Appendix D.

[6]Figure A.3 of appendix A present the same result by level of 2022 euros annualized income instead of percentiles.

have relatively high levels of pensions outside of the private sector. We also test other sample restrictions to exclude those individuals in Appendix D, with a mild effect on the observed life expectancy gradient.

For the rest of the distribution, life expectancy is almost linear in income rank. Minimum life expectancy at 55 equals 23 years and is associated to the 7^{th} percentile. Average life expectancy at the very top of the distribution is around 31 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 1.8 years. This result might be driven by lower mortality inequalities among women[7], but it is also likely to be strongly driven by the chosen income concept : labor 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. The maximum longevity difference reaches 1.8 years.

It is also worth noting that life expectancy at 55 for the first percentile of women is 32 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 43^{th} percentile and increasing above. The maximum life expectancy gap is relatively low and reaches 3 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. Secondly, life expectancy of 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 by figure 3b[8].

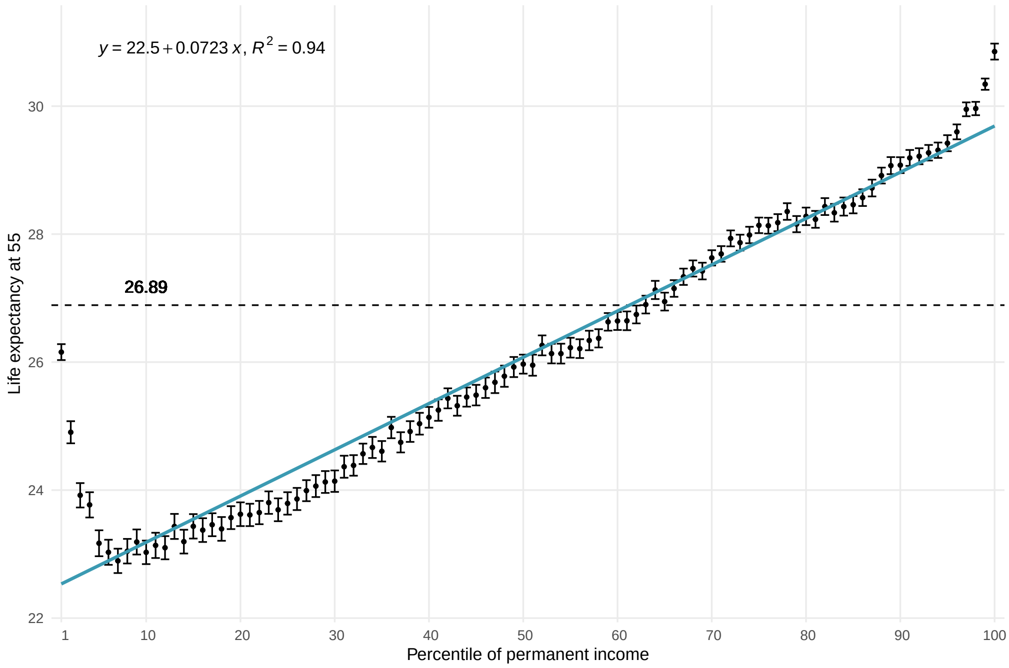
The observed gradient is the result of this composition effect between the gradient of men and that of women. This result has some important redistribution implications, visible in section 4: When abstracting from between gender variations in mortality, inequalities in life expectancy and subsequent implicit redistribution appear moderate.

[7]There is indeed less variability in age of death among women – longevity is more concentrated at some ages, as apparent in Figure A.10.

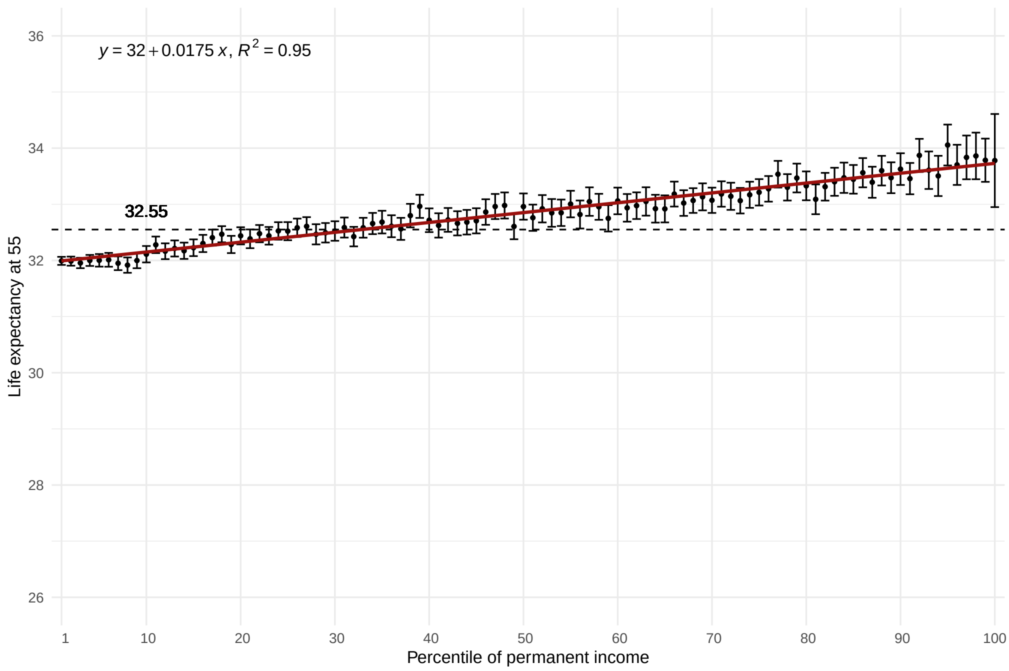
[8]Note that the share of women is actually increasing in the first few income ranks. We link this exception to our discussion on permanent income definition, for men with a low share of their career in the private sector. Those individuals are over-represented in the first income ranks.

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

(a) Men



(b) Women

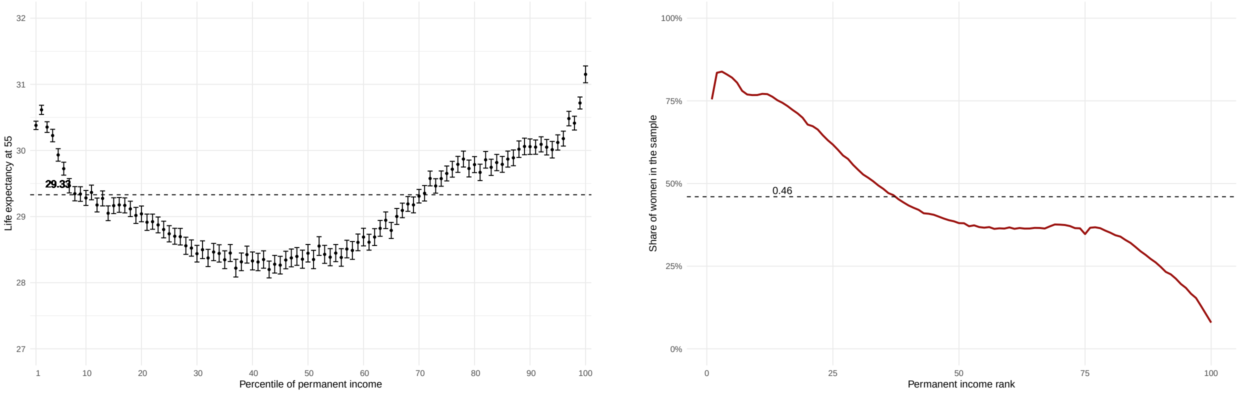


Note : This figure presents life expectancy at 55 by percentile of permanent income – average labor income between 40 and 54. Income rank is defined for the whole sample – men and women pooled. 95% confidence intervals follow the methodology of Camarda (2015). The level of income gradient is estimated 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. Life expectancy at 55 for a woman of the 30^{th} percentile is of 32.5 year
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 labor income between 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 12^{th} and 70^{th} percentiles. There are less women than average above the 34^{th} percentile. SOURCE : Agirc-Arrco data, authors' calculations

3.3 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 top and bottom vintile 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 1.8 in our study.

We attribute our comparatively lower gradients to three concurring 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.

Secondly, we use individual labor 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 a relatively large age window, thus allowing for income mobility within this timeframe. 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).

These methodological distinctions offer a plausible explanation for the lower income-related mortality gradients observed in our study, especially pronounced in the case of women.

4 Differential mortality and redistribution

We now analyse to what extent differential mortality affects redistribution in the pension system. For that purpose, we focus on the 1954 cohort, for which we have a complete view of both lifetime contributions and pension levels in the private sector: they were young enough at the creation of the second pillar scheme for executives and non-executives in 1962, and old enough to be retired as of 2022, the last observation in our data. This allows to recover both individual contributions and pension levels[9]. We restrict our analysis to affiliates who retired before this date and for whom we can reconstruct the whole flow of career contributions, namely around 314,000 individuals, of which 47% are women.

The heterogeneity in terms of life expectancy is likely to have important redistributive impacts as individuals living longer perceive more pensions than individuals with shorter lives. The magnitude of these redistributions also depends on pension formula, which operate redistributions between individuals.

4.1 Measuring redistribution using return on contributions

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

[9] Additionally, most of them claimed their pension after 2014, date from which administrative records of *Agirc-Arrco* are complete regarding pension levels. Retirement before 60 is possible under strict conditions through the "long career" scheme. We do not observe these if they die before 2014. This case is however very limited. Finally, note that individuals who died before claiming their pensions are also not included. Details on additional sample selection and treatment of mortality before retirement can be found in Appendix E.

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

where a is age, β the discount factor[10] and P_i, R_i, C_{ai}, S_{ai} are respectively the annual pension level, claiming age, annual social security contributions and survival probability function. Note that all amounts are deflated by an index of consumption prices: we cannot predict the trajectory of future pensions. However, in France, pensions are indexed on prices and therefore do not vary over age in real terms. In that sense, β should be thought of as a discounting factor on top of the price index.

Contributions are observed at each age in administrative records. Pensioners records allow us to observe claiming age and pension level at claiming. The only component we do not fully observed is mortality: on the one side, we need to account for individuals who die before claiming age, who are by definition excluded of this sample. On the other side, we do not observe age of death for individuals still alive in 2022. We circumvent these issues by leveraging our mortality results on the 1922-1964 sample. Before 55, we assume no mortality. After 55, we apply to each sub-population of the 1954 cohort the average mortality of the same sub-population, estimated on the 1922-1964 sample.

The higher the RC, the bigger the flow of perceived 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 the RC between individuals. Those are driven by three factors, namely claiming age, mortality and replacement rate – ratio of pension level over lifetime contributions. Variations in claiming age and mortality influence RC by playing on pension duration. Replacement rate can also vary between individuals based on non-contributory elements of the pension formula[11]. Decomposing the effects of these three factors on redistribution is the goal of this analysis. The methodology is detailed in the following subsection.

[10]In all that follows and unless otherwise mentioned, we use $\beta = 0.98$ as our discount factor. Other discount factors are tested in Appendix A.

[11]This includes additional rights for period of non or low employment, or advantages related to children. See appendix section C.4, for a description of the existing mechanisms in the French system.

4.2 Decomposing variations in returns on contributions

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 vintile. For each block, x, RC is obtained as :

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

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, $S_x = (S_{x8}, \ldots, S_{x68})$ 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 are 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 overall 1954 cohort as seen in sample of section 3 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 :

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

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:

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

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.

Adjusting for replacement rate In a second scenario, we further adjust for replacement rate, defined as the ratio between pensions and life-time contributions. To do so, we first define :

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

as before, we then adjust for the generation replacement rate:

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

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

Decomposition The decomposition is finally obtained using the different terms:

$$RC_K - RC_\Omega = \underbrace{RC_K - RC_K^{LE}}_{Impact\ of\ LE} + \underbrace{RC_K^{LE} - RC_K^{LE,RR}}_{Impact\ of\ RR} + \underbrace{RC_K^{LE,RR} - RC_\Omega}_{Impact\ of\ retirement\ age}$$

Note also that the result of this decomposition is not independent from the order in which we successively remove observed differentials. We therefore test the sensitivity of our results to this order in Appendix A, and find that main results are left unchanged. On top of decomposing the gender gap in RC, we also propose a within gender decomposition of the gap between income vintiles.

Interpretation We interpret variations around the population mean as an indicator of redistribution in the pension system. The RC indicator allows us to identify the biggest explicit or implicit "winners" and "losers" of redistribution, in terms of lifetime monetary transfers, compared to a benchmark of equalized return on contributions for all. This implicit equity norm has limitations, as do all inequality indicators. In particular, it is uncertain whether pension systems aim at this equalization. On the one hand, most pension systems involve a good part of contributivity. 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. In section 4, we also complement this indicator by decomposing variations in Social Security Wealth[12]. 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: one may strongly benefit from redistribution, while having a relatively low standard of living at retirement – a relatively high RC and relatively low SSW. We therefore consider RC and SSW as two complementary measures when quantifying and qualifying redistribution in the pension system.

4.3 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.

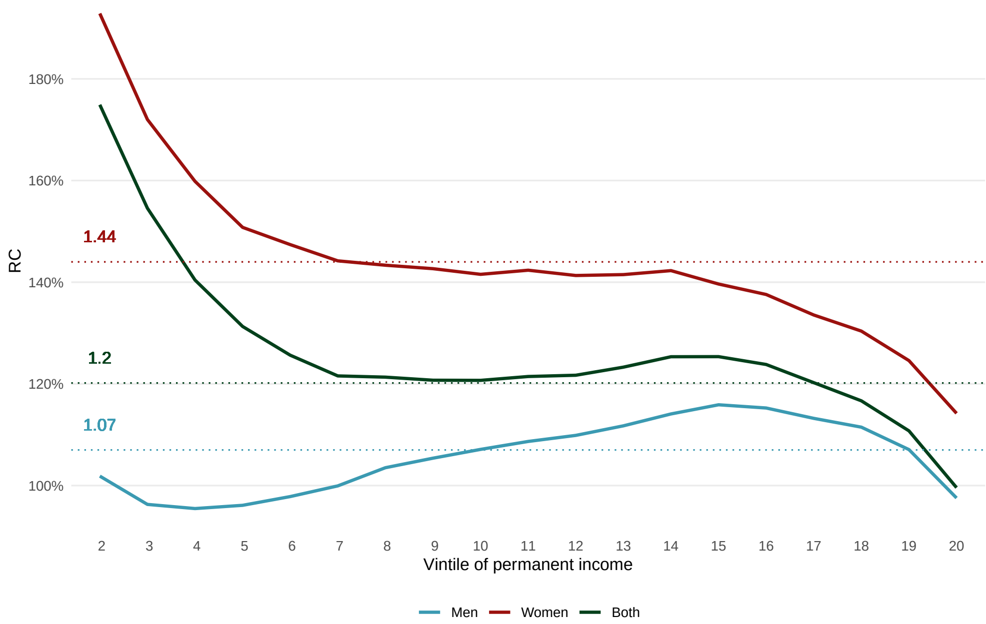
First, solidarity components of the system aim at ensuring sufficient pension levels to individuals with low levels of contribution during their careers. Provided that poor workers indeed benefit more from these non-contributory components, one expects RC to be decreasing with permanent income, translating to redistribution from rich to poor 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

[12]The methodology for this decomposition is similar to the one used for RC and can be found in Appendix E

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.

In order to assess which of the three 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 dark green line of Figure 4.[13] It appears that RC is globally decreasing with income, ranging from 205% for the first vintiles to 100% for the last ones. The overall progressivity of the system is mostly driven by the extreme parts of the income distribution, as the RC is relatively flat between the 7^{th} and 17^{th} vintiles.

Figure 4: Return on contributions for cohort 1954



NOTE : This figure presents RC by vintile of permanent income – average labor income between 40 and 54 – for all population (dark 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 vintile × gender level. It is then aggregated at the income vintile 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 144%. On average, women below the 7^{th} vintile enjoy a higher return than average.
SOURCE : Agirc-Arrco data, authors' calculations

The overall shape of the RC gradient by quantile presented in Figure 4, however, aggregates important within- and between-gender differences. When looking at the same gradient by gender, we indeed observe the following patterns: first, RC is

[13] All figures presented by income group are truncated below the 2^{nd} vintile for readibility reasons : as previously mentioned, mortality patterns for the bottom of the distribution is biased upwards (individuals with other income sources). Untruncated figures are presented in Appendix A.

higher among women on average (144% against 107%), and for every income group. This suggests that the pension system largely redistributes from men to women. Second, when looking at redistribution within gender, we observe opposite patterns among men and women. Among women, the system seems to be highly progressive. On the contrary, for men, we find an increasing RC on most of the distribution, suggesting overall regressivity. The RC gradient both genders pooled, obtained in Figure 4, is the aggregation of these results. To see this, let us recall that the share of women and men varies between the bottom and the top of the distribution – see Figure 3b. The bottom vintiles are made of a majority of women: progressivity dominates at the bottom. The top vintiles are made of a majority of men, but we observe progressivity both among men and women: progressivity also dominates at the top. In the middle of the distribution, progressivity among women is more or less compensated by regressivity among men.

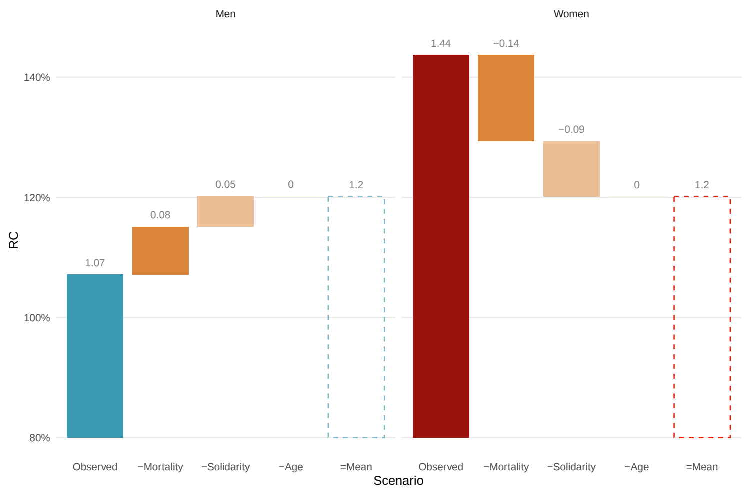
4.4 Between gender redistribution

Figure 5 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 higher return on contributions than men on average. Absent disparities in mortality, replacement rate, and retirement age between men and women, RC would indeed be 24pp lower for women and 13pp higher for men. This implies a redistribution from men towards women in a lifetime perspective.

Using the decomposition detailed at the beginning of this section, we then find that mortality differentials account for around half of these disparities. Higher longevity among women – refer to Figure 2 – leads to higher returns. Interestingly, this mortality differential has a higher impact on the gender gap in RC than differential in "solidarity" – replacement rates, despite the fact that women benefit more from solidarity of the pension formula, which aims at providing decent levels of pensions and to partially compensate for career interruptions. Finally, since women retire on average later than men – see Figure F.3 – this last factor tends to reduce the gap in RC, albeit in a rather limited way.

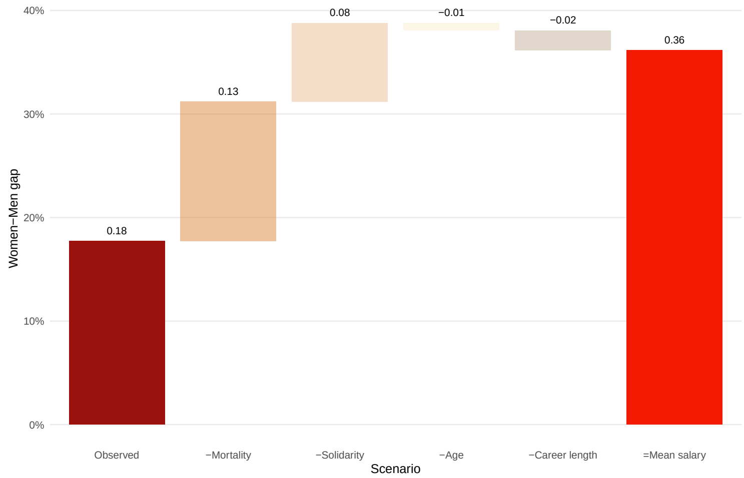
The interpretation of this result in terms of welfare is not straigthforward: women indeed benefit strongly from redistribution on average, but it does not mean that they benefit from higher standards of living at old age. It does not even mean that they benefit from higher levels of lifetime perceived pensions. We document and decompose this gender gap in lifetime perceived pensions – Social Security Wealth – in Figure 6. The gender gap in SSW is equal to 18% and is presented in the first column of the figure. Other columns present the contribution of five factors to this observed gap – inequality in mortality, replacement rate, retirement age, career duration and mean salary. We first find that without higher longevity among women, the SSW gap would be 13pp (72%) higher. Removing also the higher solidarity of the pension formula towards women, it would be around two times higher. We also find the effect of mortality to be slightly bigger than the one of solidarity. The welfare effect is however not the same: explicit solidarity of the pension formula increases standard of living in each period, while redistribution through mortality corresponds to the realization of insurance against old age poverty, to which women are more exposed on average. Finally, the remaining gap after removing mortality and solidarity differential is mainly explained by the gap in mean salary during the career.

Figure 5: 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. Three next columns successively remove differentials in mortality, solidarity and retirement age 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. Differentials in benefit from solidarity of the pension formula, and in mortality, contribute positively to this result. Differentials in retirement age contribute slightly negatively to it.
SOURCE : Agirc-Arrco data, authors' calculations

Figure 6: 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 five components. In the first column, we depict observed SSW gap. Five next columns successively remove differentials in mortality, solidarity, retirement age, 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 gap in permanent labor income. Discount rate equals 0.98.
INTERPRETATION: The average lifetime pension gap between men and women equals 18%. 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 39pp.
SOURCE : Agirc-Arrco data, authors' calculations

4.5 Within-gender redistribution

We now decompose the RC profiles observed in Figure 4, within gender and between income groups. The results are presented in Figure 7. In each graph, observed RC gradient is presented in a dark line, while average RC within gender is displayed in a light line. The gap between these lines corresponds to the lifetime redistribution operated by the pension system. For each income group, we present in three bars the contribution of each component – mortality, solidarity, retirement age– to this observed gap. We find very different explanations to the observed profiles for men and women. First, both for men and women, solidarity has a strong progressive impact on the RC profile – people at the bottom of the gradient 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 2^{nd} vintile is on average 64pp higher because of these non-contributory components – 32pp for men of the 2^{nd} vintile.

Second, differentials in retirement age profile have similar "inverse U-shape" effects on the redistributive profile among men and women : individuals at the very top and bottom of the distribution retire later than average, thus reducing RC for these income groups, all other things being equal. However, this is not the dominant effect, neither among men nor among women.

Third, we find that mortality differentials have regressive effects on the RC profile, both among 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 on most of the gradient. Among women, the mortality differentials we estimate are moderate and the order of magnitude of the effect on progressivity is low compared to solidarity. The resulting profile is therefore different: among men, regressivity dominates between the 4^{th} and 15^{th} vintile, while progressivity dominates along the whole distribution among women.

4.6 Discussion on magnitude and policy implications

Finally, let us qualify the magnitude of the observed effect of mortality differentials on redistribution. To assess it, we compute an implicit tax (subsidy) on contributions paid by short-lived (long-lived) individuals. This implicit tax corresponds to the share of contributions paid that virtually give access to no right during retirement because of longevity lower than average. It is computed by comparing observed return and return that would be observed in a scenario with homogeneous average mortality. It is therefore the tax on contributions that rationalizes pensions obtained under this counterfactual RC. Details on its computation are found in Appendix G.

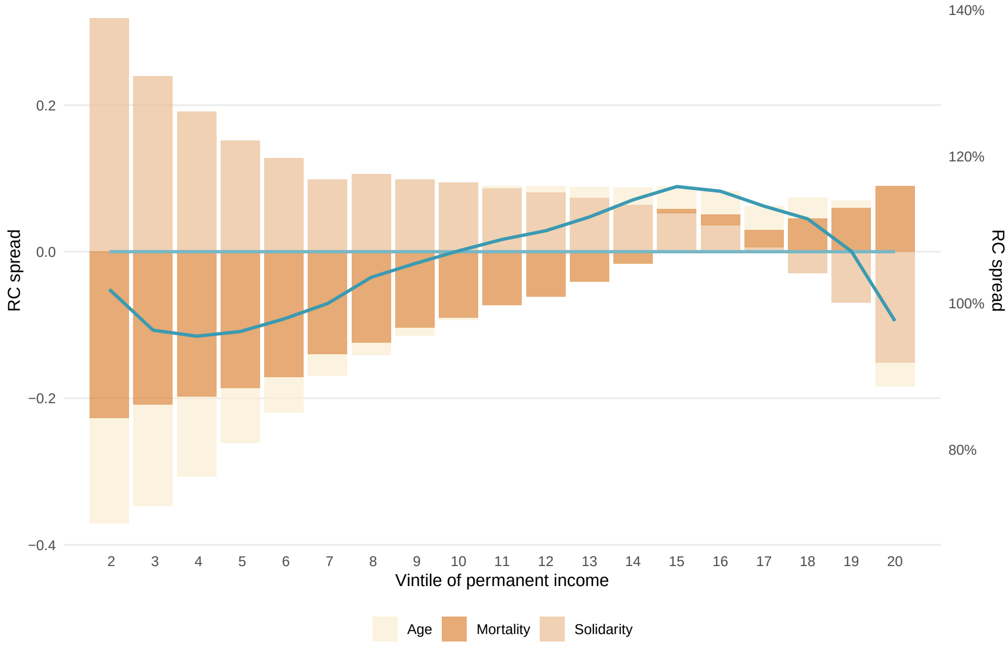
Within gender, we find an implicit tax of up to 18% of contributions at the bottom of the gradient, and an implicit subsidy of up to 10% at the top. Among women, implicit taxes and subsidies are very moderate and do not exceed 3% of lifetime contributions. Finally, between men and women, we find that mortality differentials lead to an average implicit tax of around 7% of men's contributions, that are virtually redistributed as a 11% subsidy on women's contributions.

We argue that the interpretation of these magnitudes is not the same within and between gender. Between gender, the implicit subsidy towards women through longevity is sizeable and larger than explicit redistributive features of the pension formula. But it does not come close to offsetting the observed pension gap, that is mostly driven by career inequalities – lower wages, more career interruptions, higher retirement age. Let us also note that the life expectancy gap between men and women is rather decreasing and expected to decrease in the future (Sundberg et al., 2012). In terms of policy implications for the pension system, it is therefore unsure whether the longevity gap should be a cause for concern, both for equity reasons and because its prominence should decrease.

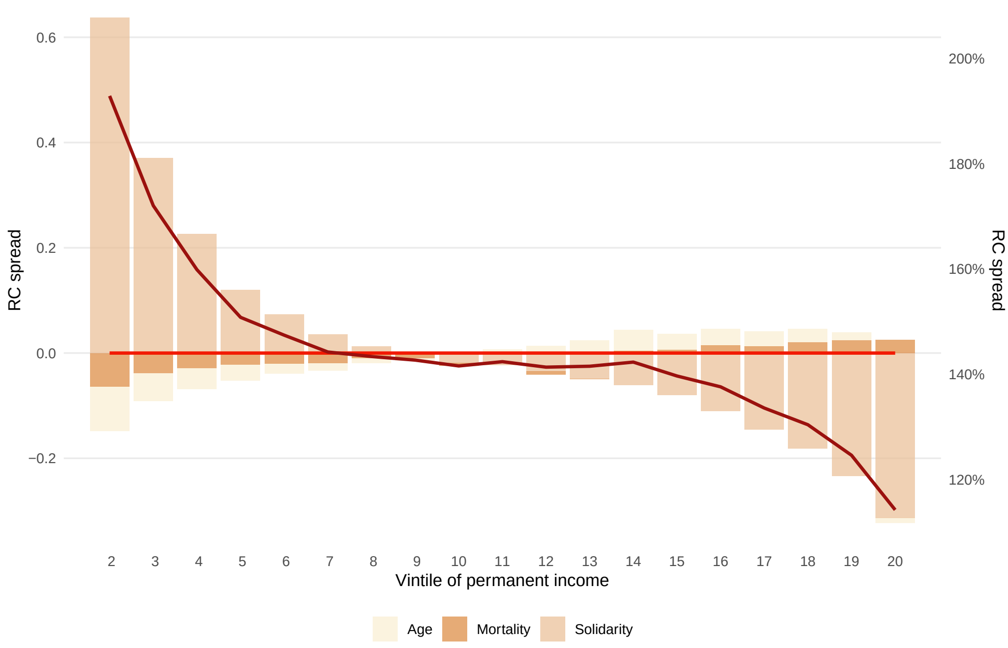
Within gender however, we argue that the large implicit tax observed among men is more likely to be cause for policy concern, for two reasons: first, it appears as a double penalty for individuals of the bottom of the distribution, since their level of pension is relatively low, and they perceive it for a shorter amount of time. Secondly, the literature on life expectancy shows that life expectancy inequalities are rather increasing in recent years (Chetty et al., 2016). The average negative effect on equity is therefore large and rather increasing. The potential impact on political acceptability of the pension system, or of pension reforms that aim at increasing retirement eligibility age, is not insignificant.

Figure 7: Decomposition of the RC gap within gender

(a) Men



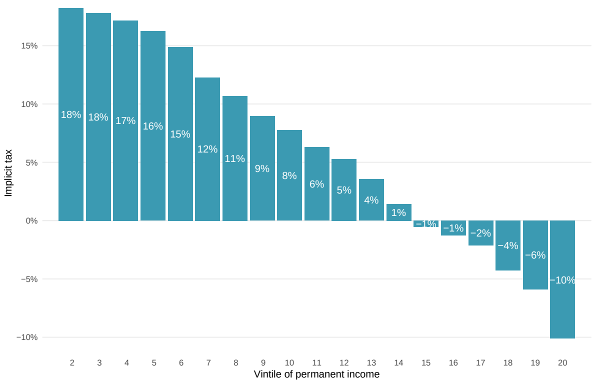
(b) Women



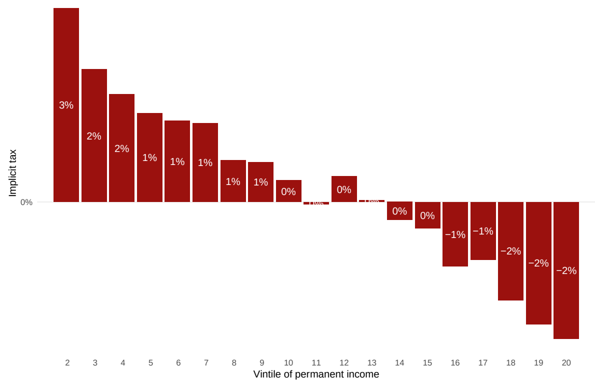
Note : This figure presents average RC by vintile of permanent income and the decomposition of between income groups, within each gender (Men in panel a, Women in panel b). Income rank is defined for the whole sample – men and women pooled. In each panel, the dark line depicts observed return on contributions by gender and income vintile. The light line depicts average RC within gender. For each income group, we decompose the gap between these two values into three components – mortality, replacement rate (solidarity), and retirement age. Discount rate equals 0.98.
Interpretation: The average return for men of the 6th vintile is lower than average : the progressive effect of the pension formula is more than offset by the effect of lower longevity.
Source : Agirc-Arrco data, authors' calculations

Figure 8: Implicit mortality tax between and within gender

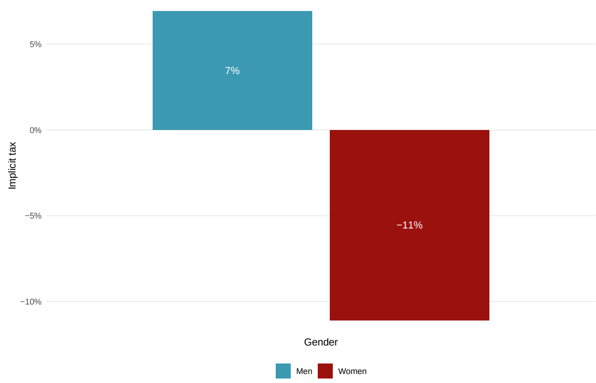
(a) Within men



(b) Within women



(c) Between gender



Note : This figure presents the level of implicit tax on contributions by vintile of permanent income within gender (panel a and b), or by gender (panel c). Income rank is defined for the whole sample – men and women pooled. Implicit tax is computed by comparing observed return to return that would be obtained absent mortality differential – within (panel a and b) or between (panel c) genders. Computation of this implicit tax is detailed at Appendix G. It corresponds to the share of paid contributions that virtually give access to no right because of lower than average longevity.
Interpretation: Compared to a scenario in which all men experience the same average mortality, contributions of men at the 8^{th} vintile are implicitly taxed at 11%, while those of men at the 20^{th} vintile are implicitly subsidized at around 10%.
Source : Agirc-Arrco data, authors' calculations

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 sizable despite this redistribution, as the observed amount of pensions perceived during lifetime is still 18% 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 1.8 years, and its observed redistributive impact is therefore also less prominent.

The implication of this hidden redistribution for policy design is not straightforward. 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 realization 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 high (Baurin, 2021). 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, thus hiding 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

Alvaredo, F., Atkinson, A. B., Blanchet, T., Chancel, L., Bauluz, L., Fisher-Post, M., ... others (2020). *Distributional national accounts guidelines, methods and concepts used in the world inequality database* (Unpublished doctoral dissertation). PSE (Paris School of economics).

André, M., Germain, J.-M., & Sicsic, M. (2023). 'do i get my money back?': A broader approach to inequality and redistribution in france with a monetary valuation of public services. *Document de travail Insee*(2023-07).

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

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).

Bozio, A., Garbinti, B., Goupille-Lebret, J., Guillot, M., & Piketty, T. (2024, 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

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

Dherbécourt, C., Didier, M., & Lefebvre, G. (2023). *La protection sociale en france et en europe en 2022 - résultats des comptes de la protection sociale - Édition 2023.* Panoramas de la DREES. Retrieved from https://drees.solidarites-sante.gouv.fr/publications-communique-de-presse-documents-de-reference/panoramas-de-la-drees/la-protection-0 (Published on 14 December 2023)

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.

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

Id 010752373. (2023). Retrieved 2023-04-04, from https://www.insee.fr/en/statistiques/serie/010752373 (Accessed: 2023-04-04)

Id 010752374. (2023). Retrieved 2023-04-04, from https://www.insee.fr/en/statistiques/serie/010752374 (Accessed: 2023-04-04)

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.

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.

Sundberg, L., Agahi, N., Fritzell, J., & Fors, S. (2012). Why is the gender gap in life expectancy decreasing? the impact of age- and cause-specific mortality in sweden 1997–2014. *Solidarité santé. Etudes statistiques*(33).

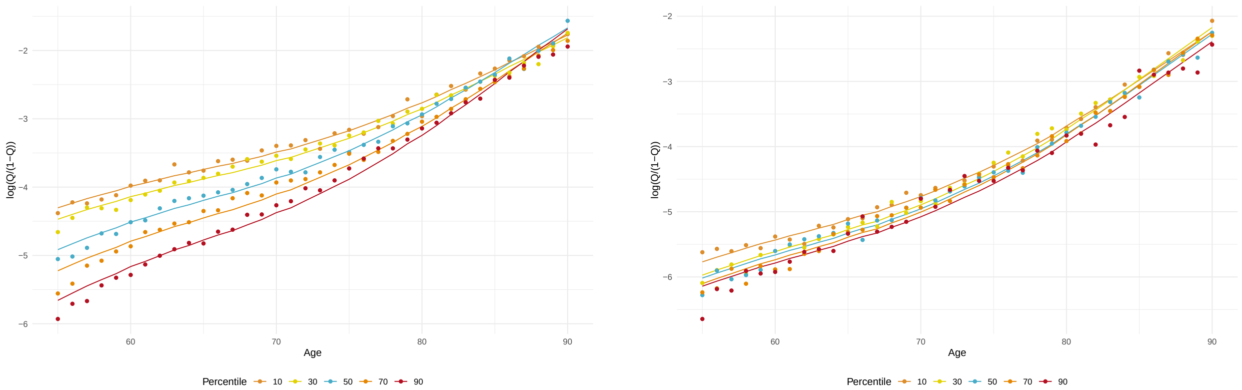
Tréguier, J. (2021). Retirement: Do women and men make the same decisions? *Revue économique*, *72*(6), 881-928. Retrieved from https://www.cairn-int.info/journal-revue-economique-2021-6-page-881.htm doi: 10.3917/reco.726.0881.

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

A Additional Tables and Figures

Figure A.1: Mortality rates by income group

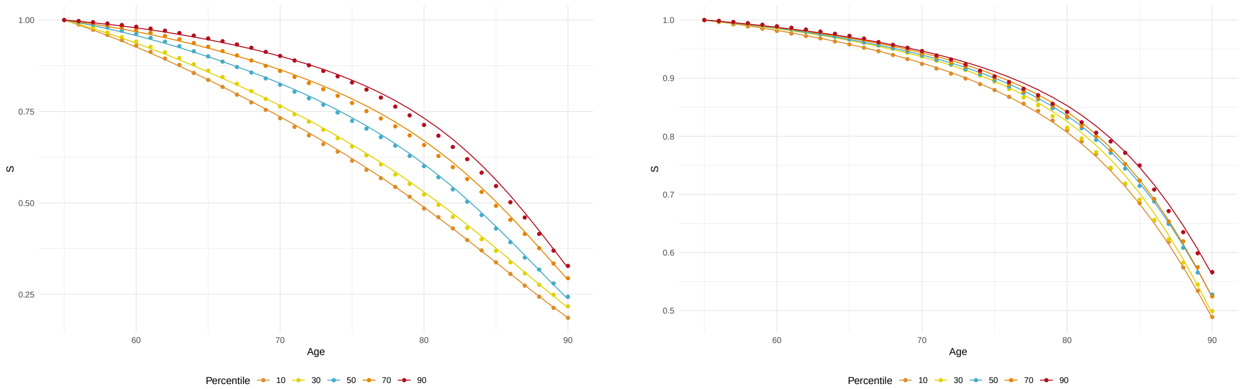
(a) Men (b) Women



NOTE : This figure presents estimated – line – and observed – dots – mortality rates by age for the 10^{th}, 30^{th}, 50^{th}, 70^{th}, 90^{th} percentiles of permanent income, by gender. More precisely, it presents the logistic transformation of these rates. For each percentile, mortality is estimated out of observed mortality using a logistic model between 55 and 90. Income rank is defined for the whole sample – men and women pooled.
INTERPRETATION: A man of the 10^{th} percentile has an estimated mortality rate at 60 of around 1.8% (logistic transformation equal to -4).
SOURCE : Agirc-Arrco data, authors' calculations

Figure A.2: Survival rates by income group

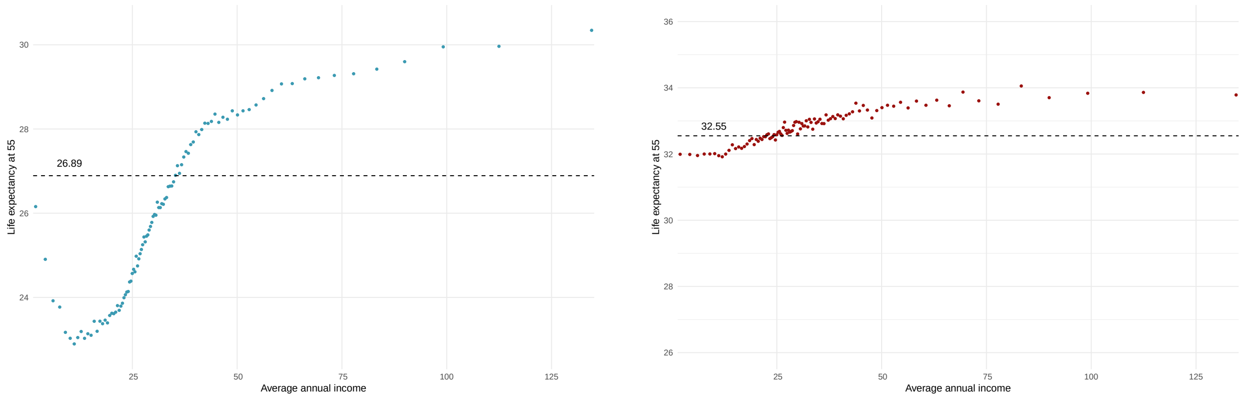
(a) Men (b) Women



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 logistic 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

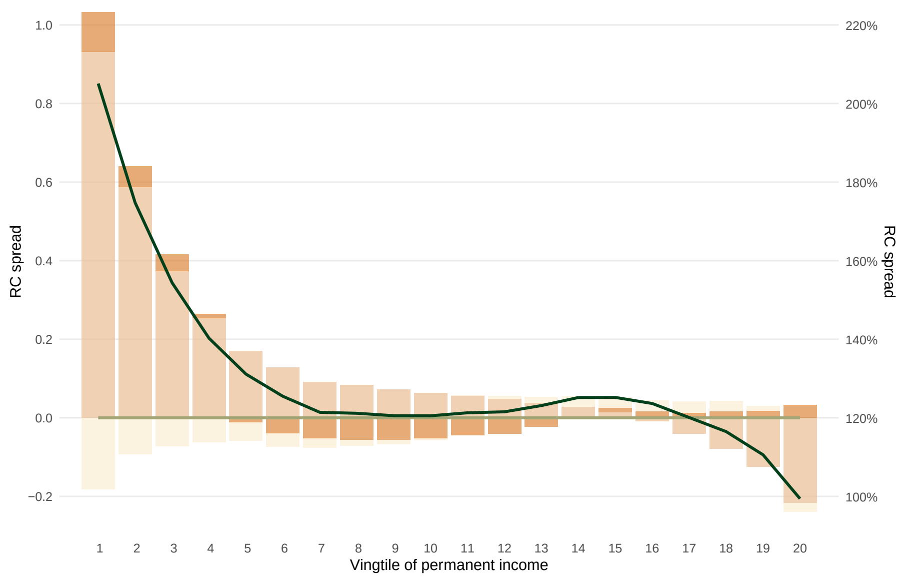
Figure A.3: Life expectancy at age 55 by income level

(a) Men (b) Women



Note : This figure presents life expectancy at 55 by level of permanent income – average labor 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 annualized terms and in 2022 €.
Interpretation: Men with average labor income of 50 k€ between 40 and 54 have a life expectancy at 55 of around 28 years.
Source : Agirc-Arrco data, authors' calculations

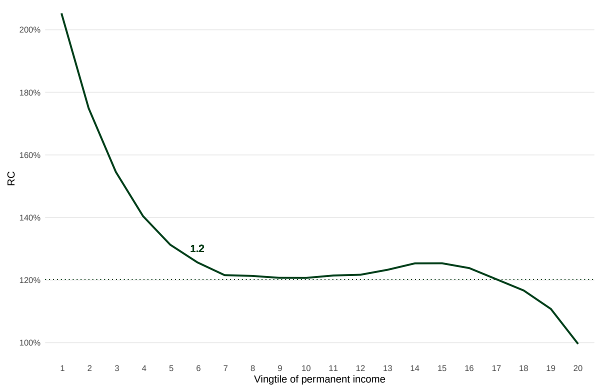
Figure A.4: Decomposition of the RC gap – all genders pooled



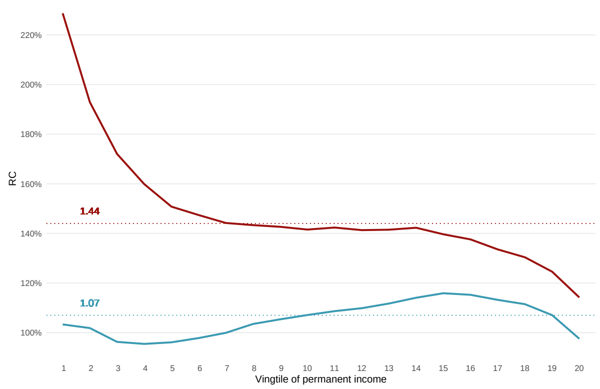
Note : This figure presents average RC by vintile of permanent income and the decomposition of between income groups, for men and women ppoled. Income rank is defined for the whole sample – men and women pooled. The dark line depicts observed return on contributions by income vintile. The light line depicts average RC for the whole cohort. For each income group, we decompose the gap between these two values into three components – mortality, replacement rate (solidarity), and retirement age. Discount rate equals 0.98.
Interpretation: The average return for individuals of the 6th vintile is slightly higher than average : The regressive effect of lower longevity than average is more than offset by the progressivity of the pension formula.
Source : Agirc-Arrco data, authors' calculations

Figure A.5: RC figures – non truncated

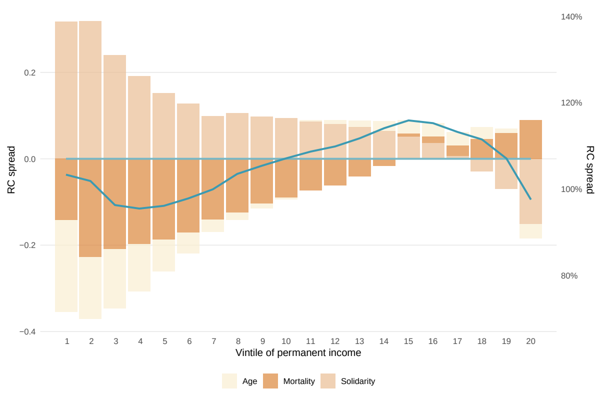
(a) RC by income group



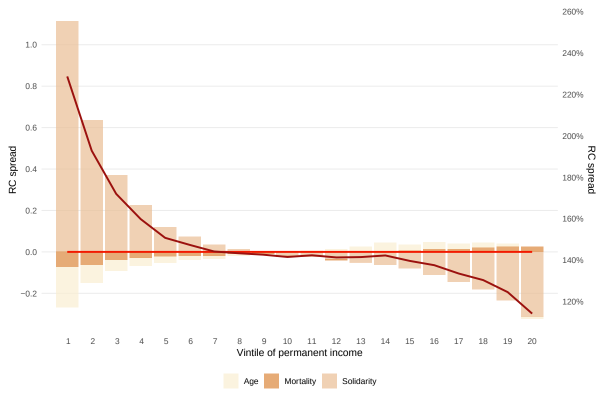
(b) RC by income group and gender



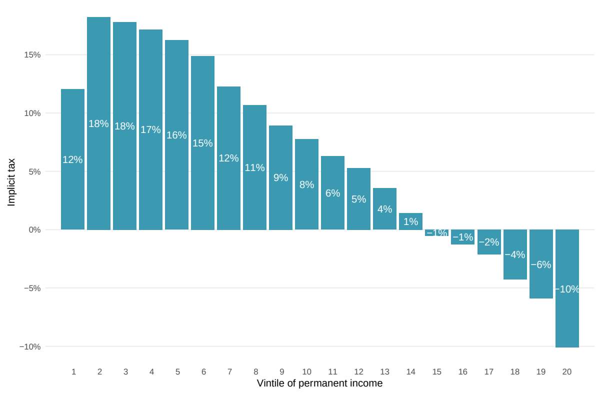
(c) RC gap within gender – Men



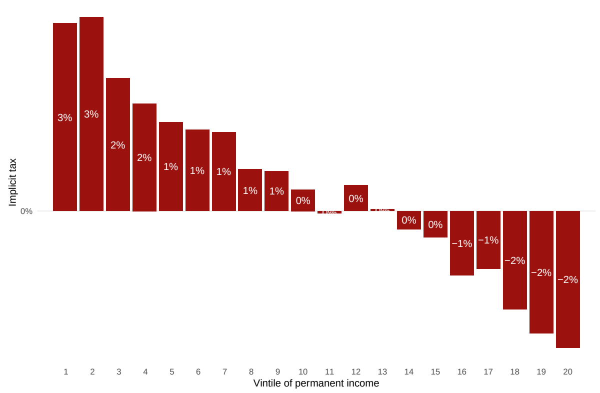
(d) RC gap within gender – Women



(e) Implicit mortality tax within men



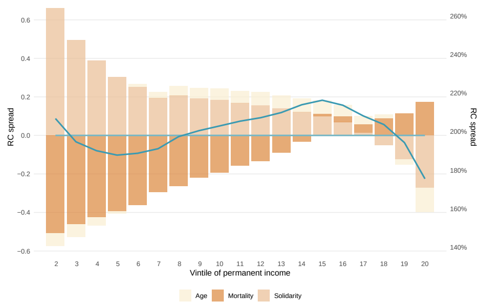
(f) Implicit mortality tax within women



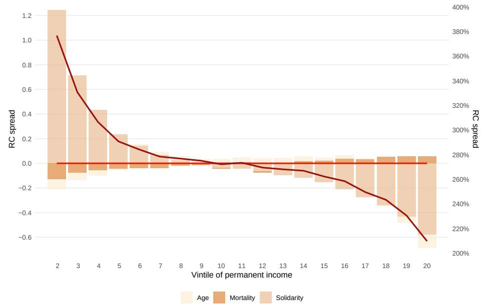
NOTE : This figure presents results of section 4 by income group by including all permanent income groups – including vintiles 1 and 2. Corresponding truncated figures are respectively Figures 4, 7a, 7b, 8a, 8b . Discount factor equals 0.98.
SOURCE : Agirc-Arrco data, authors' calculations

Figure A.6: Decomposition of the RC gap within gender – by discount factor

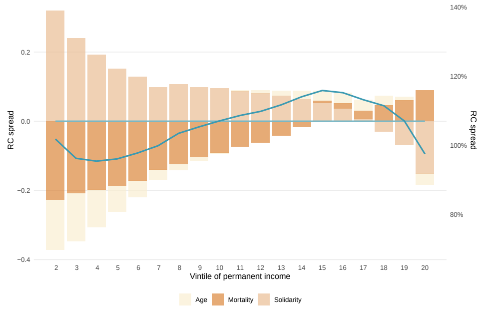
(a) Men – $\beta = 1$



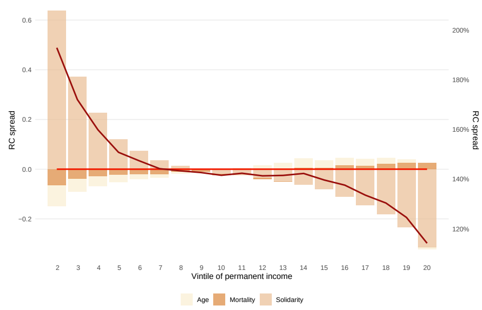
(b) Women – $\beta = 1$



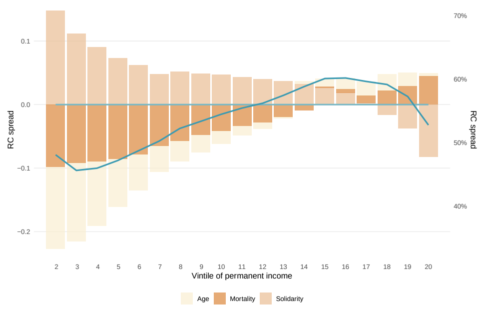
(c) Men – $\beta = 0.98$



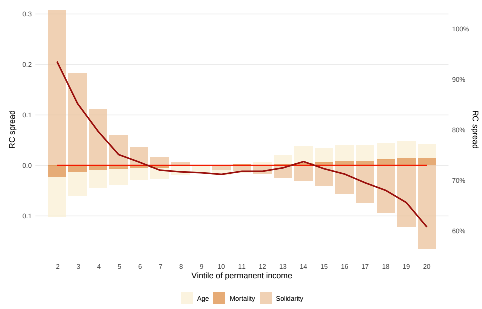
(d) Women – $\beta = 0.98$



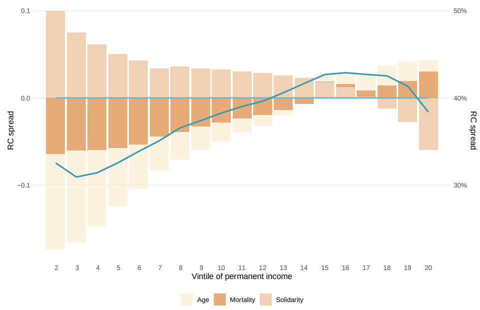
(e) Men – $\beta = 0.96$



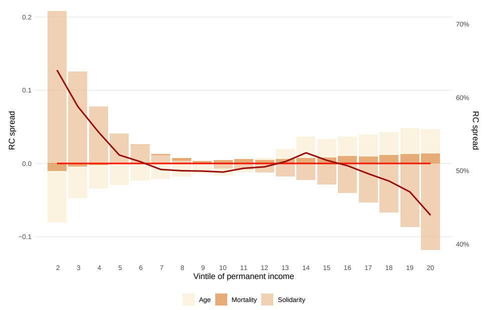
(f) Women – $\beta = 0.96$



(g) Men – $\beta = 0.95$



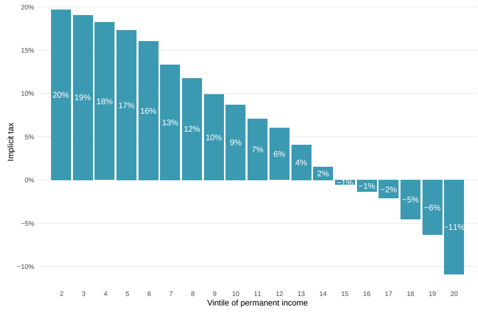
(h) Women – $\beta = 0.95$



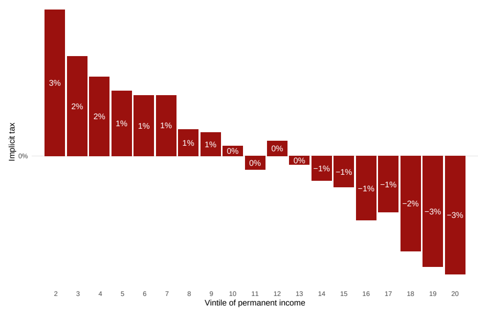
Note : This figure presents average RC by vintile of permanent income and the decomposition of between income groups, within each gender (Men in panel a, Women in panel b), using four different values for the discount factor – 1, 0.98, 0.96 and 0.95. Income rank is defined for the whole sample – men and women pooled. In each panel, the dark line depicts observed return on contributions by gender and income vintile. The light line depicts average RC within gender. For each income group, we decompose the gap between these two values into three components – mortality, replacement rate (solidarity), and retirement age. Discount rate equals 0.98.
Interpretation: The higher the discount factor, the lower the computed RC for a given income group and gender – past contributions have a higher weight than future pensions.
Source : Agirc-Arrco data, authors' calculations

Figure A.7: Implicit mortality tax within gender – by discount factor

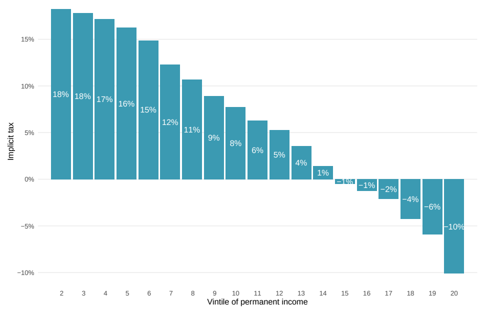
(a) Men – $\beta = 1$



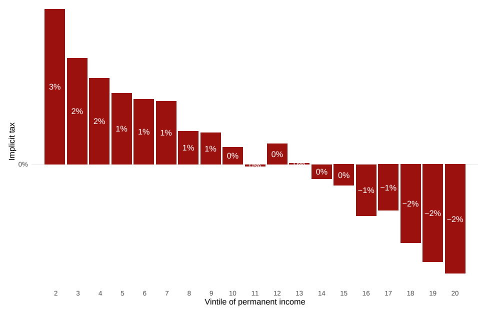
(b) Women – $\beta = 1$



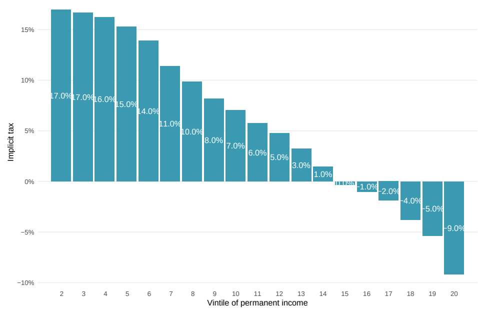
(c) Men – $\beta = 0.98$



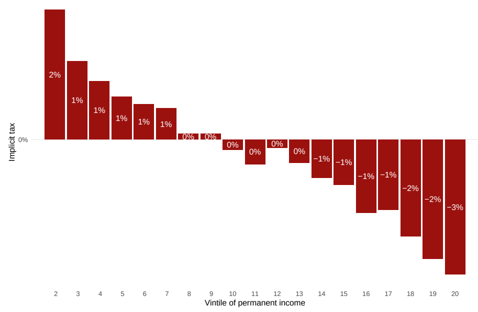
(d) Women – $\beta = 0.98$



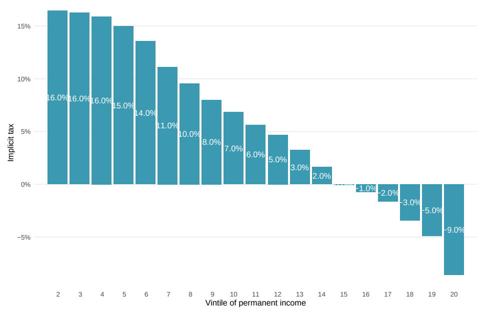
(e) Men – $\beta = 0.96$



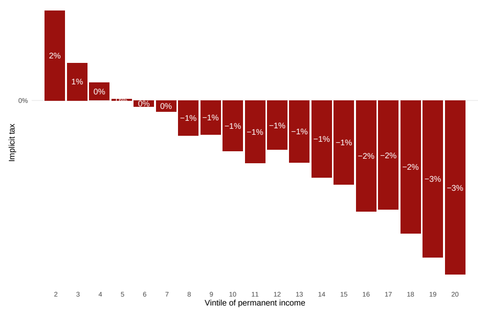
(f) Women – $\beta = 0.96$



(g) Men – $\beta = 0.95$



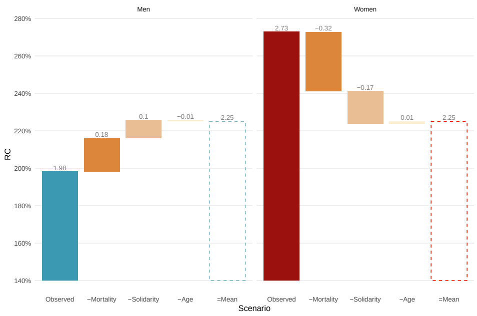
(h) Women – $\beta = 0.95$



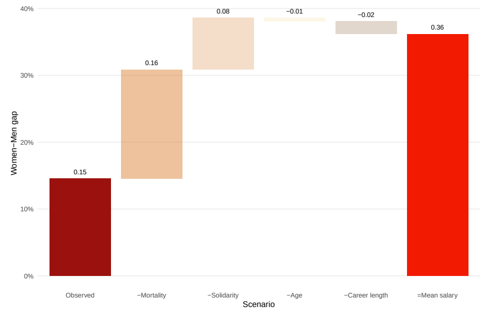
NOTE : This figure presents the level of implicit tax on contributions by vintile of permanent income within gender (panel a and b), or by gender (panel c), for four different values of the discount factor – 1, 0.98, 0.96, 0.95. Income rank is defined for the whole sample – men and women pooled. Implicit tax is computed by comparing observed return to return that would be obtained absent mortality differential – within (panel a and b) or between (panel c) genders. Computation of this implicit tax is detailed at Appendix G. It corresponds to the share of paid contributions that virtually give access to no right because of lower than average longevity.
INTERPRETATION: The higher the discount factor, the lower the effect of mortality on the RC in absolute terms – old age periods receive less weight. Therefore, the higher the discount factor, the lower the implicit tax rate on contributions due to lower longevity.
SOURCE : Agirc-Arrco data, authors' calculations

Figure A.8: Decomposition of the RC and SSW gap between gender

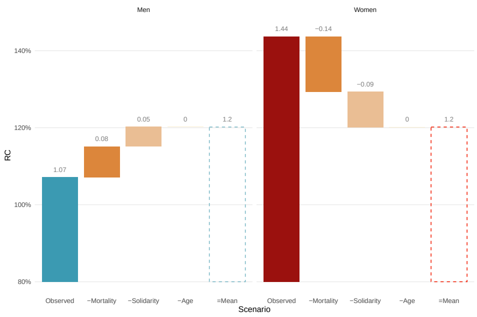
(a) RC – $\beta = 1$



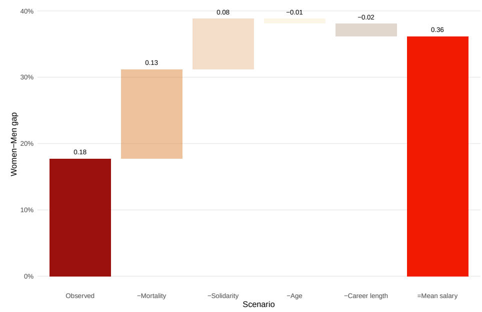
(b) SSW – $\beta = 1$



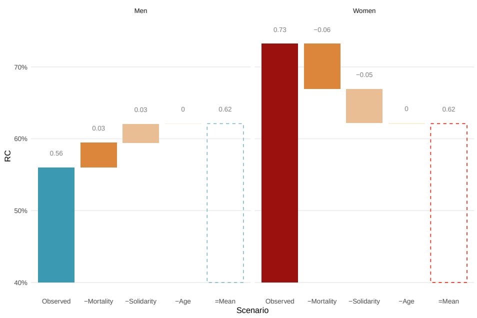
(c) RC – $\beta = 0.98$



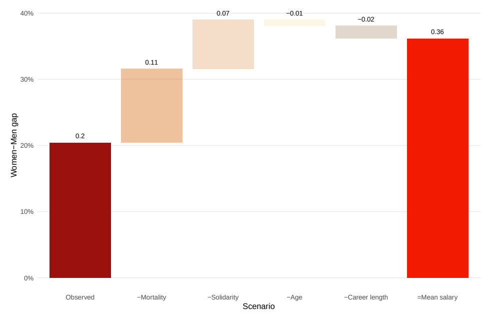
(d) SSW – $\beta = 0.98$



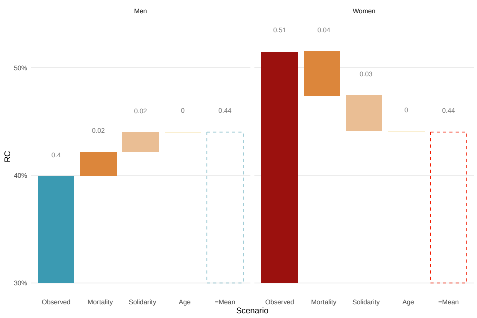
(e) RC – $\beta = 0.96$



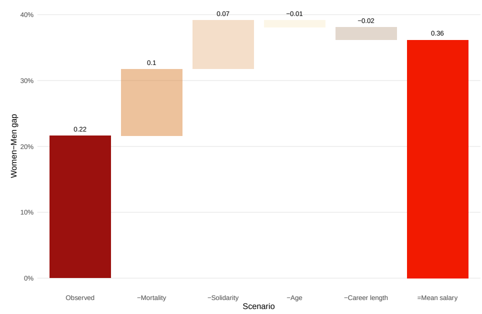
(f) SSW – $\beta = 0.96$



(g) RC – $\beta = 0.95$

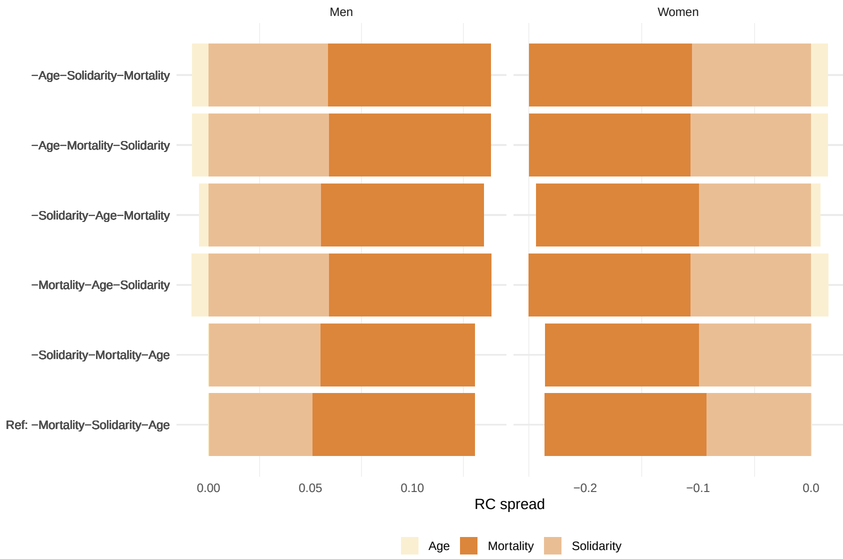


(h) SSW – $\beta = 0.95$



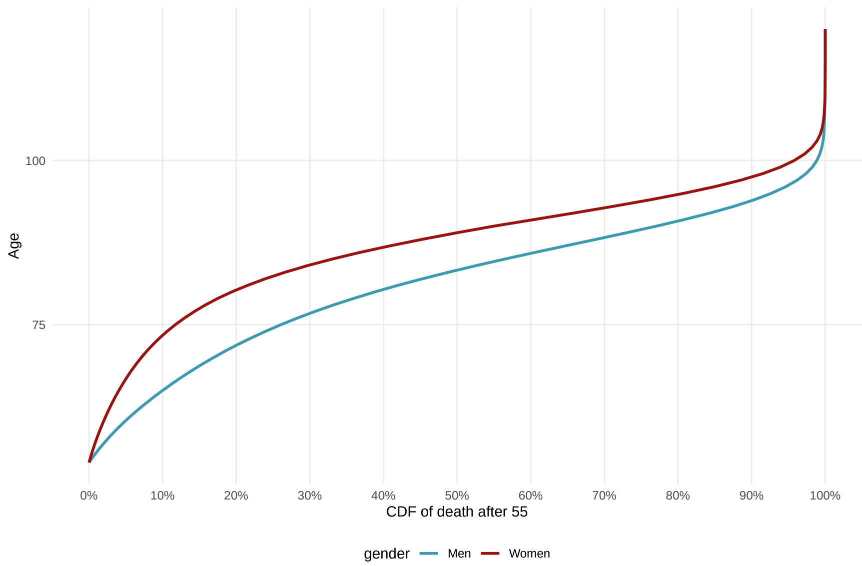
NOTE : This figure presents average RC by gender and its decomposition – left panels – and decomposition of the lifetime pension gap between gender – right panels – using four different values of the discount factor – 1, 0.98, 0.96, 0.95. Left panels are computed in the same way as Figure 5. Right panels are computed in the same way as Figure 6.
INTERPRETATION: The higher the discount factor, the lower the relative impact of mortality on the RC and SSW gaps – old age periods receive less weight.
SOURCE : Agirc-Arrco data, authors' calculations

Figure A.9: Order sensitivity in the RC gap decomposition



Note : This figure presents the respective contributions of differentials in retirement age, mortality and solidarity – replacement rate – in the observed gap in RC between men and women, for different orders of decomposition. Each line corresponds to an order of decomposition – three factors, six combinations : the first one from the bottom is the benchmark scenario developed in section 4 and depicted in Figure 5. Discount factor equals 0.98.
Interpretation: In all six methodologies, mortality is the main contributor to the observed gap in RC between men and women.
Source : Agirc-Arrco data, authors' calculations

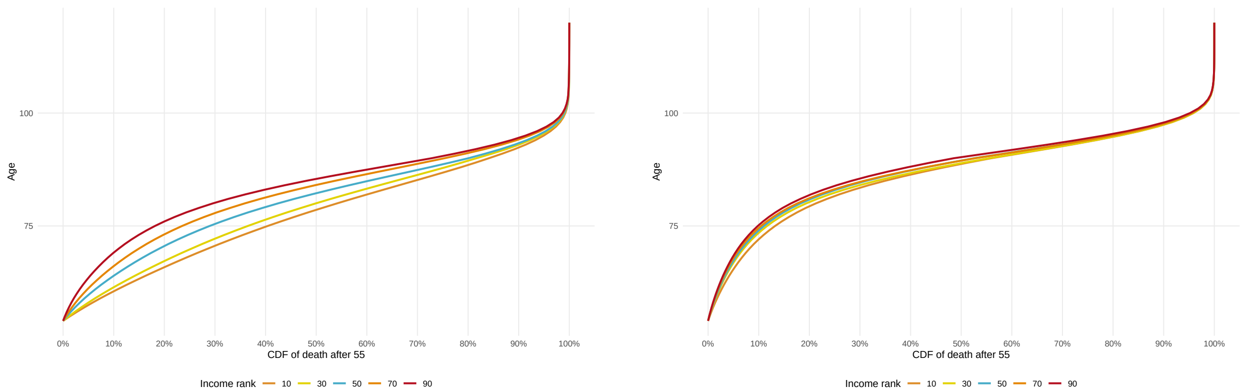
Figure A.10: Cumulative distribution function of longevity by gender



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 3.
Interpretation: 12% of women alive at 55 die before reaching 75, against 25% of men.
Source : Agirc-Arrco data, authors' calculations

Figure A.11: 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 3.
INTERPRETATION: Among men of the 90^{th} percentile, 30% die before the age of 80.
SOURCE : Agirc-Arrco data, authors' calculations

B Details on data sources

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.

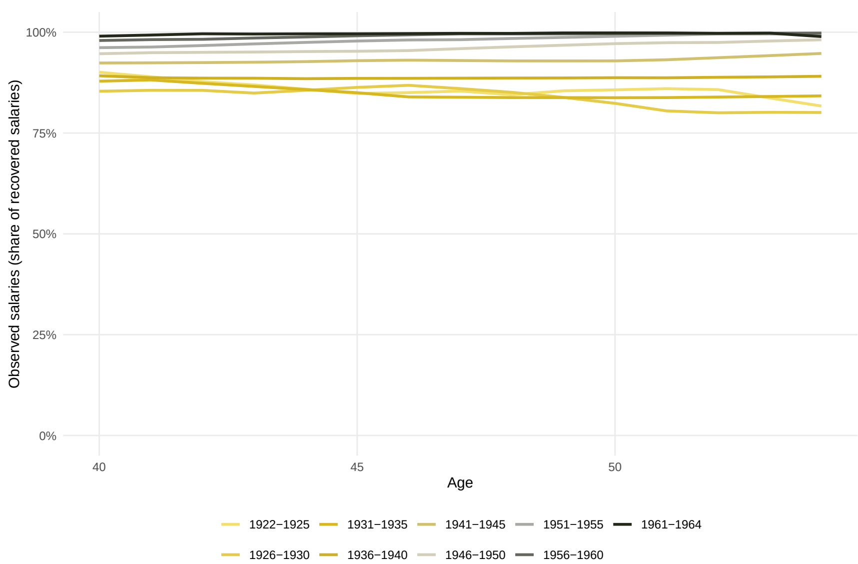
Third, *Agirc-Arrco* records provide information on pensioners in different databases. We first use *Bases allocataires* – pensioners database, one per year between 2009 and 2022, that contains information on total points validated at claiming, by regime (*Agirc*, *Arrco*, unified *Agirc-Arrco*) and 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 Appendix C. From *RCIV* records, we recover total periods validated by pensioners of the regime up to 2022. These periods are differenciated 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 perceived 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* (*ID 010752373*, 2023, *ID 010752374*, 2023).

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 labor income out of incomplete career data. We use earnings data when available. However, for older cohorts, earnings data is more incomplete than points data, as apparent in Figure B.2. When points are available and earnings are not, we reconstruct salaries from points, using the price of points and estimates of individual contribution rates. The detail of this methodology is given is Appendix C.3. A related difficulty is that, by definition, we only observe earnings for periods spent within one of the *Agirc-Arrco*. Labor income earned in other sectors – *e.g.* the public sector – are therefore not observed. For these two reasons, in our permanent income computation, we normalize the sum of labor 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 which we assume observe income is more or less representative of earnings on the whole career. In section 3, 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 who retired and for which we have data on pensions, we do find evidence of this : individuals of the first five vintiles of 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 **??**

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



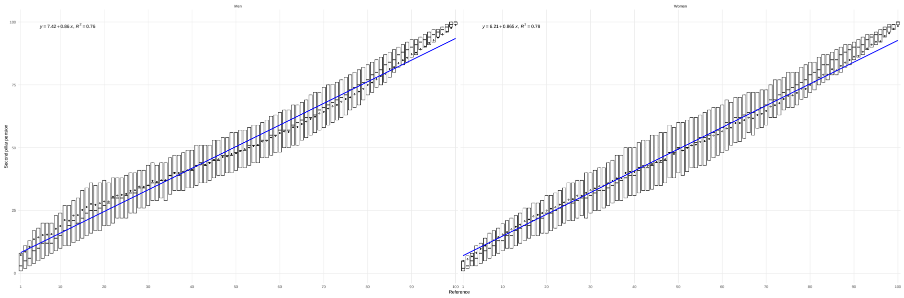
Note : This figure presents total salaries from earnings database over total recovered earnings by age and cohort group. . Total recovered earnings equal observed earnings when available and earnings recovered from points when not. These are obtained by multiplying points by purchase price and dividing by individual contribution rates, following methodology detailed in C.3.
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 retain earnings top-coded at 4SSC for computation of permanent income. However in section 4, we retain all contributions up to 8SSC. In Figure B.3, we display, for cohort 1954, the distribution of observed earnings (brown, top-coded at 8SSC after 1988) and the same 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.

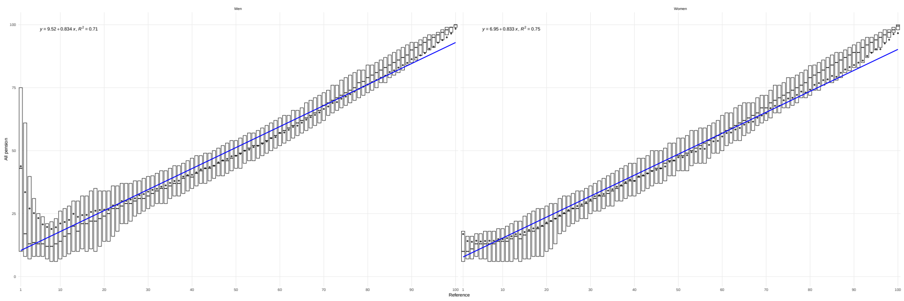
Finally, let us note that even though data on points is less incomplete than

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

(a) Agirc-Arrco pension



(b) Total pension

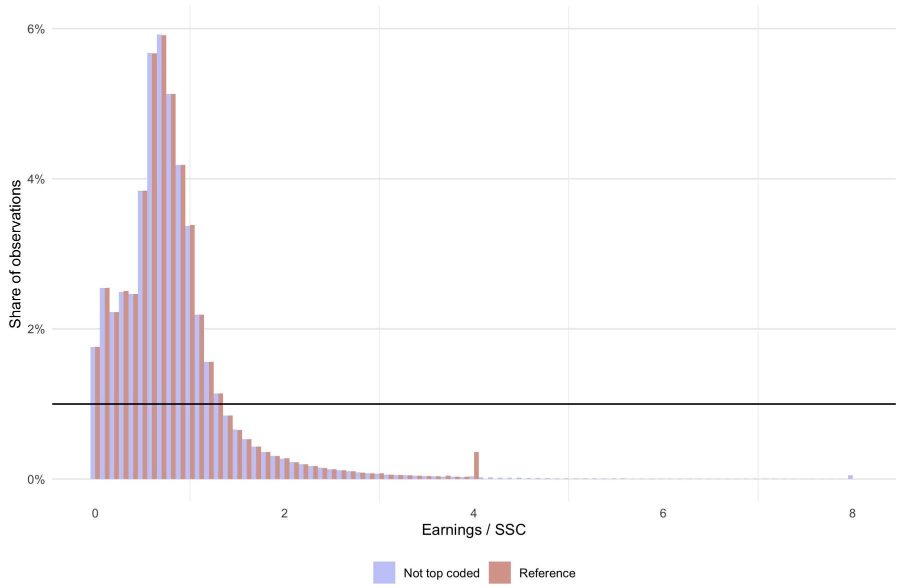


NOTE : This figure presents the observed distribution statistics of pension quantiles, by percentile of permanent labor income. For each quantile of permanent income, horizontal tick, diamond and box respectively represent median, average and interquartile range of pension percentile. Sample is made of individuals of section 3 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 40^th percentile of total pensions.
SOURCE : Agirc-Arrco data

earnings data for older cohorts, it is still rather incomplete, as apparent in Figure B.4.Indeed, total points in the career database do not come close to total points claimed at retirement for older cohorts. In the computation of permanent income, our main strategy is to normalize 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.5 : for older cohorts, earnings of the end of the career are over-represented 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 D, with low impact on our results.

Missing career data is especially cause for concern in our lifetime redistribution analysis. Indeed, we need to make sure that we recover full contribution history for our sample. The first step is to focus on a recent cohort, namely 1954, for which career data is relatively complete. We then predict missing earnings on periods

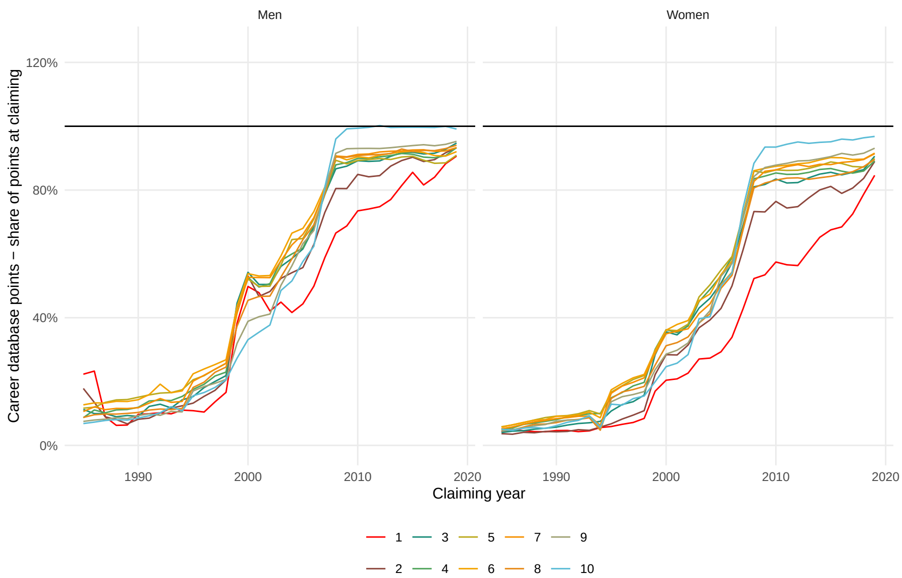
Figure B.3: Distribution of earnings for the 1954 cohort



Note : This figure presents the distribution of observed earnings, normalized 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

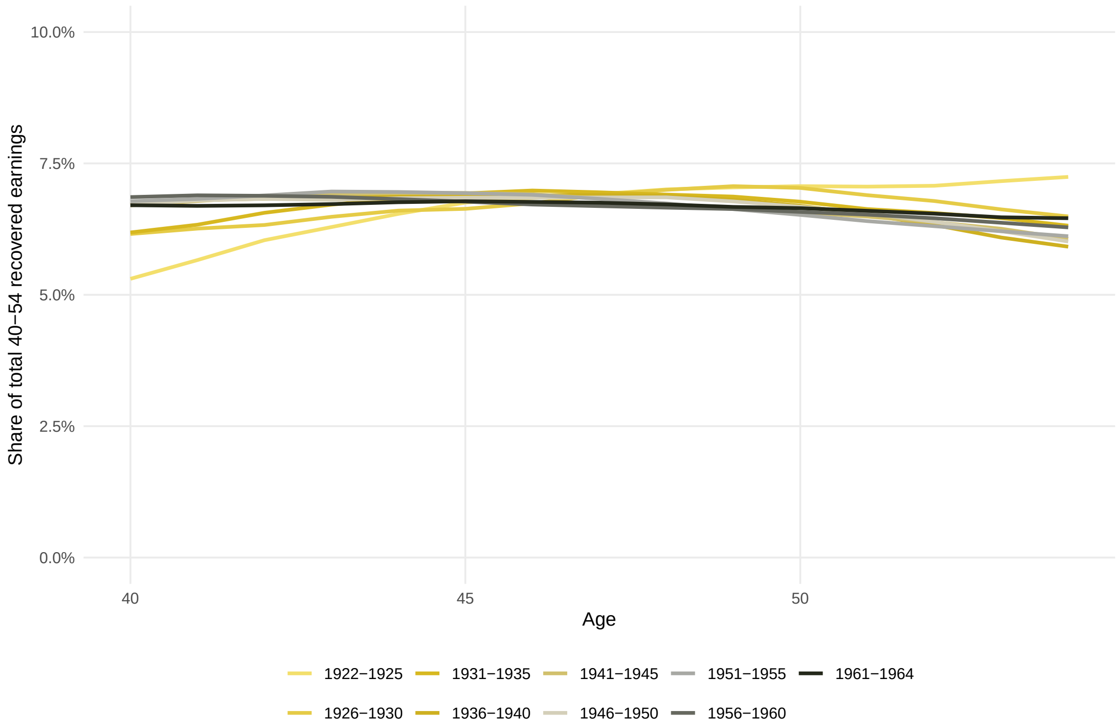
where contributed quarters are observed, but no points and earnings. We therefore compute contributions out of observed point and earnings data when available, and from predicted earnings when not. Finally, we adjust the obtained contributions to ensure consistency with point and quarter information at claiming. The precise methodology is given in Appendix E.

Figure B.4: 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 3, 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 9^{th} decile who retired in 2000, 60% of total points at claiming are missing in career records.
Source : Agirc-Arrco data

Figure B.5: 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 normalized by an index of mean salary – *SMPT*. Sample is the one used in section 3.
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

C Pension formulas in the private sector

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.

C.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:

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

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:

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

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.

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 oragnisation adminstrating 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 :

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

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[14] 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 like during the great recession.

C.2 Contribution base

The contribution base of the two pillar differ. Contribution base are defined relative to the *Plafond de la Sécurité Sociale* – Social Security Ceiling (hereafter SSC) which

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

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.[15] 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.

C.3 Computation of annual earnings for permanent income computation

Annual earnings are one of the main variables of interest of this study. We mentioned in Appendix B that data on earnings is incomplete. We also mentioned that data on points for earlier periods of the *Agirc-Arrco* records all less incomplete. When earnings are available, we use them. When only points are available, we compute labor income from accumulated points. The last term of equation (C.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 thresh-

[15] 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).

old 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 C.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 mobilize 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 :

$$W_{irt} = \sum_{b} \frac{Points_{irbt} PP_{rbt}}{\tau_{rbt}} \tag{C.4}$$

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, 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 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 C.2.

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

Bracket :	Arrco 1		Arrco 2		Agirc 1		Agirc 2	
Fixed-effects								
id	Yes	No	Yes	No	Yes	No	Yes	No
cohort	No	Yes	No	Yes	No	Yes	No	Yes
sex	No	Yes	No	Yes	No	Yes	No	Yes
year	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
quantile	Yes	Yes	Yes	Yes	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
Adj. R^2	0.624156	0.343142	0.747389	0.31929	0.824093	0.681939	0.667017	0.444719

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

In Appendix D, we test an alternative methodology where we compute earnings using historical average contribution rates observed for each bracket of the second pillar scheme, an information by the French Retirement Guidance Council (COR, 2009). We find that the life expectancy gradient is left unchanged. We however argue that this methodology is likely to introduce noise in our measure of permanent income and verify it in Figure C.1 : on observations where we observe both earnings and points, the distribution of recovered earnings using predicted individual rates is closer to observed earnings than the one obtained using mean contribution rates.

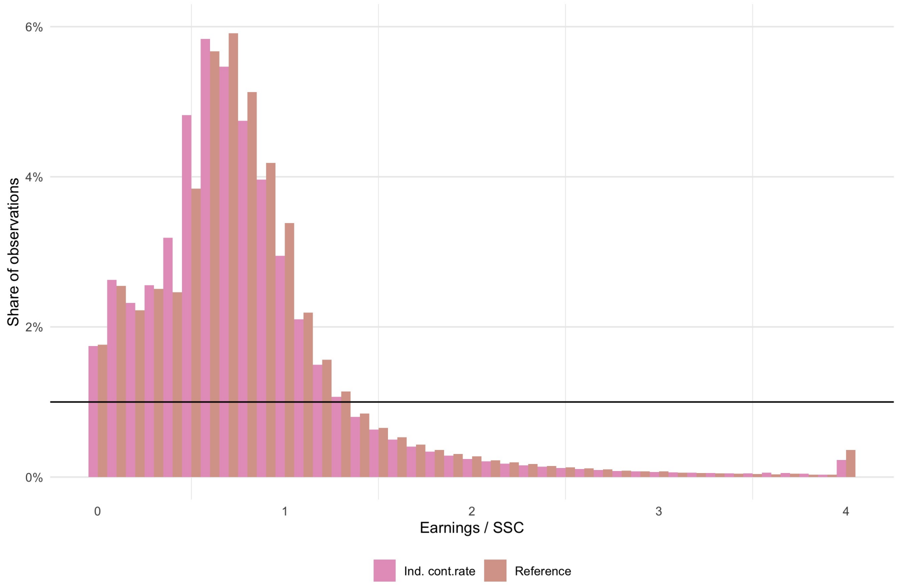
C.4 non-contributory components

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.

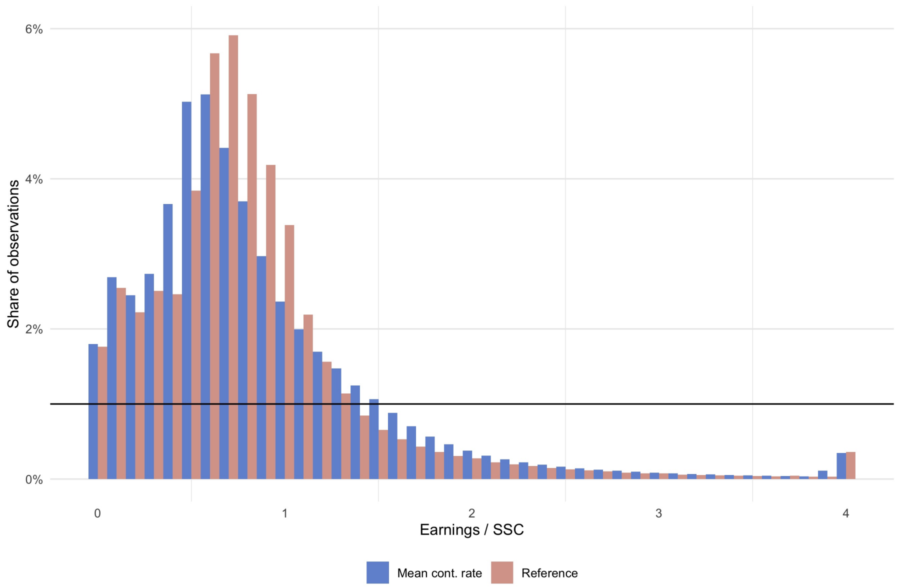
Implicit non-contributory components Given that it is based on points, the pension formula makes a clear link between contributions and pension. This is not the case of the first pillar based on annuity. In particular the way insurance duration is computed generates additional pension rights for people 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

Figure C.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, normalized by the value of the Social Security Ceiling – SSC and top-coded at 4SSC. Observed earnings are presented in brown. In C.1a, earnings recovered from estimated individual rates are displayed in pink. In C.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

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.[16] 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.

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

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

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[17] has important redistributive implications : workers with more increasing earnings benefit more from the system than others.

C.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 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.

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

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.

D Life expectancy gradient – Methodology and sensitivity analysis

Computation of life expectancy 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 count of survivors as of January 1^{st} of the considered year, and the number of deaths during the year. Age is the one that would be reached during the year. Observed mortality rates between 55 and 90 are smoothed using a logistic model between 55 and 90, using mortality rates of the overall French population as a reference point:

$$\begin{cases} \ln\left(\frac{Q_{xa}}{1-Q_{xa}}\right) & = \alpha'_x + \beta'_x a \\ \ln\left(\frac{Qinsee_a}{1-Qinsee_a}\right) & = \alpha'' + \beta'' a \end{cases}$$

Where $Qinsee$ denotes observed mortality coefficients in the overall population as measured by the National Institute for Statistics (*Insee*), and Q the underlying mortality coefficients we try and estimate, as opposed to the observed $\hat{Q}$. These two linear relationships with age boil down to:

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

Note that when sub-population x denotes individuals of the same gender, we use gender-specific mortality rates in the overall population $Qinsee_s$ in the above formula.

To compute survival functions and life expectancy at 55, we use predicted mortality rates from this regression, denoted $\tilde{Q}$, from age 55 to 90. Above 90, we use sex-specific income-independent mortality rates from the overall population. By definition, survival at age 54 is equal to 1 and for all years 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 surviving probabilities at all subsequent ages ($S_{x54} = 1$ and $S_{x121} \approx 0$):

$$
\begin{aligned}
E_{x55} &= \frac{55(S_{x55} - S_{x56}) + 56(S_{x56} - S_{x57}) + \ldots + 120(S_{x120} - S_{x121})}{S_{x54}} \\
&= \sum_{a=55}^{120} S_{xa}
\end{aligned}
$$

Sensitivity analysis Figure D.1 shows alternative measures of income gradient by gender, obtained through variations on the methodology used in section 3. Our main estimate of life expectancy gradient appears mostly robust to these variations, that can be classified in four types.

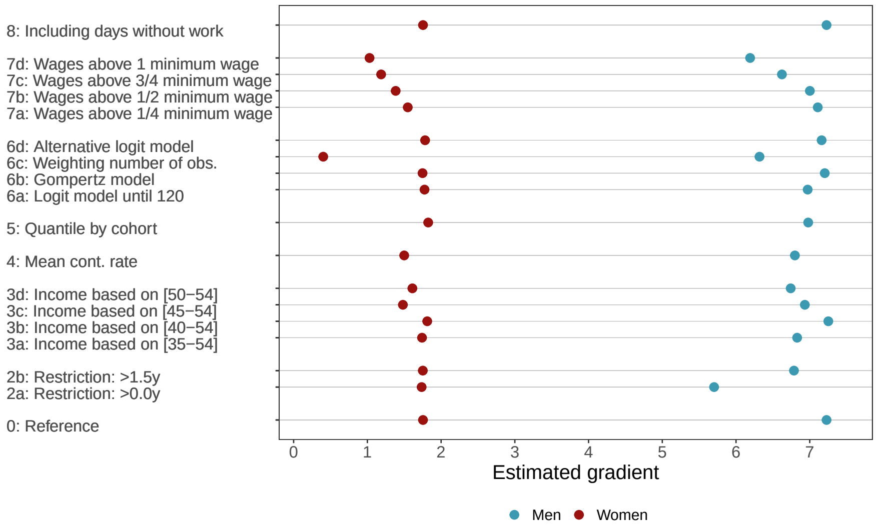
In terms of sample selection, we test different restrictions of contribution duration in the *Agirc-Arrco* regime between 40 and 54, from merely one day to 3 years, our main estimate. The higher the threshold, the more individuals with short and non-representative careers in the private sector are excluded. The income gradient tends to increase slightly with this threshold, among men only. This is consistent with what we observed in Appendix B : individuals with other sources of individdual labor income at the bottom of the gradient are mostly men. We also try to exclude observations where observed labor 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 labor, which is why we do not exclude them in our main estimate. The higher the restriction, the lower the observed longevity gradient.

Alternative income group measures are also tested. At the numerator of our permanent income measure, we try recovering missing earnings from points and mean contribution rates, instead of predicted individual rates, and find little impact on the gradient. At the denominator of our permanent income measure we try normalizing labor income by years spent at least partially spent in the *Agirc-Arrco regime*, instead of solely days of employment. The rationale is to control for individuals with high wage but long periods of work interruption, which represent negative income shocks. Failing to take into account these periods might create a positive bias in our income measure for these individuals. However, this alternative measure also has a limit : since we do not observe income outside of work, it implicitly considers null income for periods of interruption. The effect on the gradient is relatively low. Finally, we try and define quantiles of permanent income by cohort, instead of relying on the income distribution of cohort 1964. 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 labor 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 decrease in gradient of males using this methodology.

We then play on the lower bound of the age window from 35 to 50. For this analysis, 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 wider the age range, the more income mobility is taken into account. On the other hand, there is a sample selection effect. People working between 50 and 54, in lieu of being unemployed or inactive, might be relatively wealthier and healthier than those working between 35 and 54. Figure D.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 (repectively in yellow and orange on the graph). Narrower older age windows select a more homogeneous population, thus reducing the life expectancy gradient. This sensitivity analysis has little effect on observed gradient, suggesting that both effects offset each other.

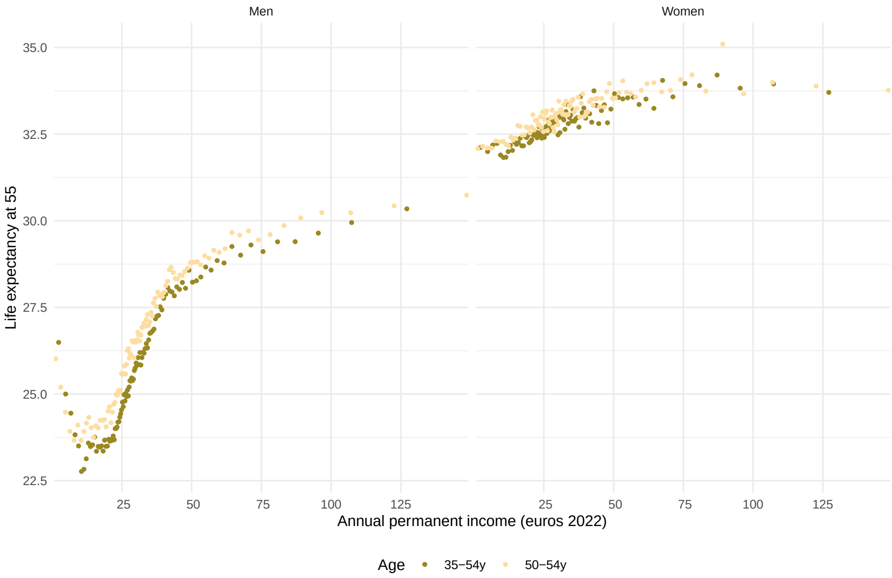
Finally, we test alternative mortality models, for life expectancy estimation. First, differential mortality is extrapolated until 120, instead of 90 in the main model. The effect of adding differential mortality between 91 and 120 on the gradient is limited. Second, we swap the logistic model for a Gompertz model between 55 and 90, with a limited impact on the gradient. This model is for example the one used by Chetty et al. (2016) between 40 and 89 and implies that log mortality rates evolve linearly with age ($\ln(Q_{csy}) = \alpha'_{cs} + \beta'_{cs}y$). The third alternative controls for for variations in the headcount across ages, in the model's regression, by weighting the number of observations at each age. Because of selection issues for older cohorts and mortality, headcount is indeed fewer at old ages. This alternative thus gives more weight to differential mortality at young ages, and infers a steep reduction in life expectancy gradient (-0.9 for men and -1.3 for women). Finally, we test an alternative specification of the logit model, when the logistic transformation of mortality coefficient is regressed on age, and not the logistic transformation of mortality coefficient for the overall population. The effect on the observed gradient is close to null.

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



Note : This figure presents life expectancy gradients by gender for different methodologies and by gender. Life expectancy gradient is defined as the estimated coefficient from the regression of life expectancy at 55 on percentile of labor permanent income. The reference point is the methodology described in 3: Average gross labor 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 day in the regime is around 1.5 years lower.
Source : Agirc-Arrco data, authors' calculations

Figure D.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 annualized 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

E RC – Details on methodology

Sample selection Compared to sample of section 3, our analysis of lifetime redistribution implies three additional sample constraints. These are detailed in Table E.1. First, we need to focus on a cohort that is old enough to have reached retirement, and young enough so that available data on careers in the first and second pillars are close to complete. We select cohort 1954, that was 8 at the creation at the second pillar scheme for non-executives, and that is at least 67 in 2022, the last year of observation and the normal retirement age in France – age at which all workers become eligible for a full rate pension. As expected, the impact on total sample size is large – around 4% of the previous sample is selected. This is however not cause for concern in itself. Indeed, focusing on one cohort is even desirable for our analysis, as it enables us to abstract from cohort effects in observing RC differentials.

Secondly, we need to observe pension levels and claiming age in both the first pillar and second pillar. Individuals who did not yet claim one of these two pensions, who died before claiming it or died before 2014 – first year of observations of all pensions in the pensioners database – are not observed. It represents 19% of the sample. In the next paragraph, we explain how we still account for the risk of mortality before retirement, by doing some assumptions on comparability of surviving and dead individuals, within income group and gender.

Thirdly, we need to account comprehensively for lifetime contributions of all remaining individuals of the sample, to compute lifetime returns on contributions. For most years of career of cohort 1954, we do observe salaries and accumulated points, that enable us to compute contributions. For some observations, however, we observe a contributed quarter, which is an evidence that contributions were paid, without observing related points and earnings. We need to make some assumptions and predictions on the actual level of earnings to compute contributions, based on years where earnings are observed. For a small share finally, we are not able to make such imputations – no year with observed earnings. We exclude individuals concerned by such a missing observation – around 5% of the sample. The precise methodology of these predictions is found in the next paragraph.

Finally, we end up with a sample of around 314,000 individuals for our lifetime redistribution analysis, of which 46.6% are women, a share very close to the one in sample of section 3 – 46.5% – and quite close to the initial sample of Agirc-Arrco affiliates – 46.2%.

Table E.1: Additional restrictions and sample size

Restrictions	Gender	Nb of obs. (k)	% previous step	% of Women
Life expectancy sample	M	5,247	-	
	F	4,562	-	46.5
1954 cohort	M	214	4.1	
	F	195	4.8	47.8
Observed pension	M	175	82	
	F	158	81	47.4
Full contribution history	M	168	96	
	F	146	92	46.6

NOTE : This table presents sample size for successive sample restrictions, by gender, for redistribution analysis in section 4. The initial sample is the one used in section 3. Each subsequent line represents an additional sample restriction.
SOURCE : Agirc-Arrco data

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 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 C.

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[18]. 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 :

[18]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

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 E.1 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 vintiles and genders, except for a few exceptions – bottom of the distribution among men and women, upper-middle quantiles among women.

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 vintile – that do not appear on the formula above for readibility reasons – and an individual fixed effect as our predictors of earnings :

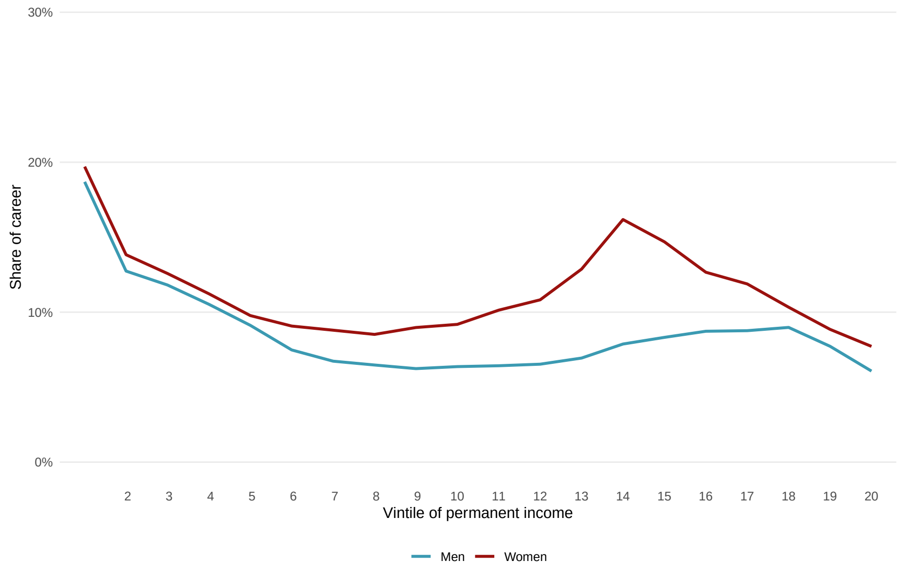
$$log(salary_{ia}) = \alpha_i + \beta_1 a + \beta_2 a^2 + \beta_3 Qua_{ia} + \beta_4 CumQua_{ia} + \beta_5 CumQua_{ia}^2 + \epsilon_{ia}$$

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 E.2 presents the results of these fixed effect regressions. Note first that 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.

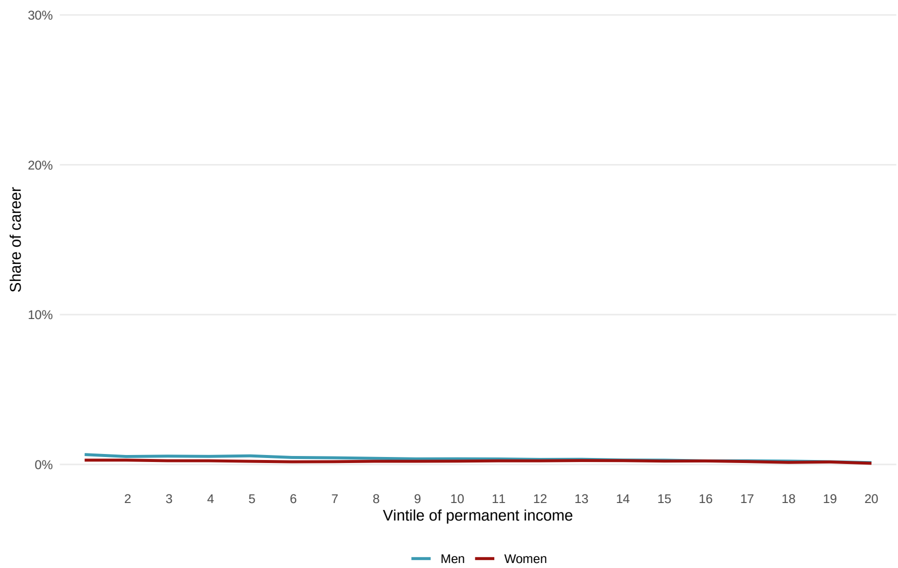
Computation and decomposition of SSW To decompose the effects of differentials in mortality, replacement rate, retirement age and career on Social Security Wealth between gender, we follow a methodology that is close to the one developed

Figure E.1: 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 vintile 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 same one as in section 4.
INTERPRETATION: On average, around 10% of *CNAV* work periods among men of the 7^{th} vintile are spent outside of the *Agirc-Arrco* regime or missing in the *Agirc-Arrco* database.
SOURCE : Agirc-Arrco data

for RC in section 4. 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 vintile.

For each block x, observed SSW is:

Table E.2: Fixed effect regression – log earnings

Dependent Variable :	log(sal)			
Model :	(1)	(2)	(3)	(4)
Variables				
age	0.641***	−0.057***	0.398***	0.024***
age^2	−0.010***	0.001***	−0.006***	0.000+
Qua	0.476***	0.450***	0.373***	0.388***
CumQua	0.005***	0.014***	0.026***	0.005***
$CumQua^2$	0.000*	0.000***	0.000***	0.000***
$age \times c$	0.000***	0.002***	0.002***	0.001***
$age^2 \times c$	0.000***	0.000***	0.000**	0.000***
$Qua \times c$	0.000***	−0.003***	−0.001***	−0.002***
$CumQua \times c$	0.000***	0.000***	0.000***	0.000***
$CumQua^2 \times c$	0.000***	0.000***	0.000***	0.000**
Fixed-effects				
id	Yes	Yes	Yes	Yes
Fit statistics				
Observations	2 935 752	4 681 563	2 077 276	4 069 662
R^2	0.786	0.622	0.734	0.653
Within R^2	0.725	0.159	0.573	0.224

*Signif. Codes : ***: 0.001, **: 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.

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

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 overall 1954 cohort as seen in sample of section 3 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 five counterfactuals successively remove differentials in mortality, solidarity, retirement age, career length, mean salary. SSW is respectively denoted SSW^{LE},

$SSW^{LE,RR}$, $SSW^{LE,RR,CA}$,$SSW^{LE,RR,CA,CL}$ and $SSW^{LE,RR,CA,CL,SA}$ in each of these counterfactuals. Formula of the three first counterfactuals are given, for a given population $K \in \{M, F\}$, by :

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

where LTC still denotes lifetime contribution and α_Ω denotes the common replacement rate for scenarios without solidarity differential. Finally, note that we get $SSW_K^{LE,RR,CA,CL}$ by computing average permanent income within population K. Indeed, one can see that the remaining differential in $SSW_K^{LE,RR,CA}$ 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 :

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

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

F Measures of career disparity between gender

Compared to men and on average, women suffer from two main types of inequalities in career trajectory. First, they are over-represented in low-wage quantiles, and under-represented in high-wage quantiles, as apparent in Figure 3.

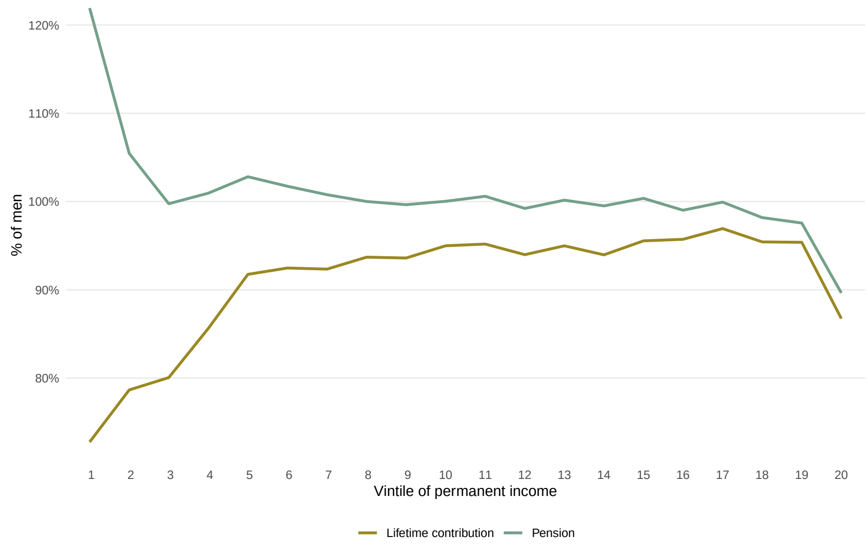
Second, for a given average wage level, women are more exposed to periods of work interruption. They thus cumulate less lifetime contributions and fewer periods of work. Figure F.1 shows that this phenomenon is true on the whole distribution of average labor income, and especially at the bottom of the distribution[19]. At the bottom of the distribution, one can speak of a "double penalty": Women are over-represented in low-wage jobs, where they are also more subject to career interruptions.

Career inequalities are partially compensated through non-contributory elements of the pension formula, that come compensate incomplete careers. The effect is double : first, since some periods of career interruptions, especially on account of children, are accounted for in retirement eligibility criteria, eligibility age for a full rate pension comes sooner. Figure F.3 displays average claiming age by income group and gender. We see that average retirement age is higher among women on almost all the distribution, despite this partial compensation of career interruptions. The second effect is an increase in the full rate pension level, relatively to lifetime contributions. We observe the first effect when looking at relative pensions of women in Figure F.1. On the whole distribution, women perceive higher pension levels, relatively to lifetime contributions, especially at the bottom of the distribution[20]. Another way of looking at this result is by comparing payback duration at Figure F.2 – that is pension duration that equalizes the flow of lifetime pensions and lifetime contributions. Since the replacement rate of women at claiming is higher than the one of men, payback duration is also lower on the whole distribution, and especially at the bottom.

[19]We also observe it in the last quantile, but it could be heavily driven by very top wages of the top quantile, which are mostly men.

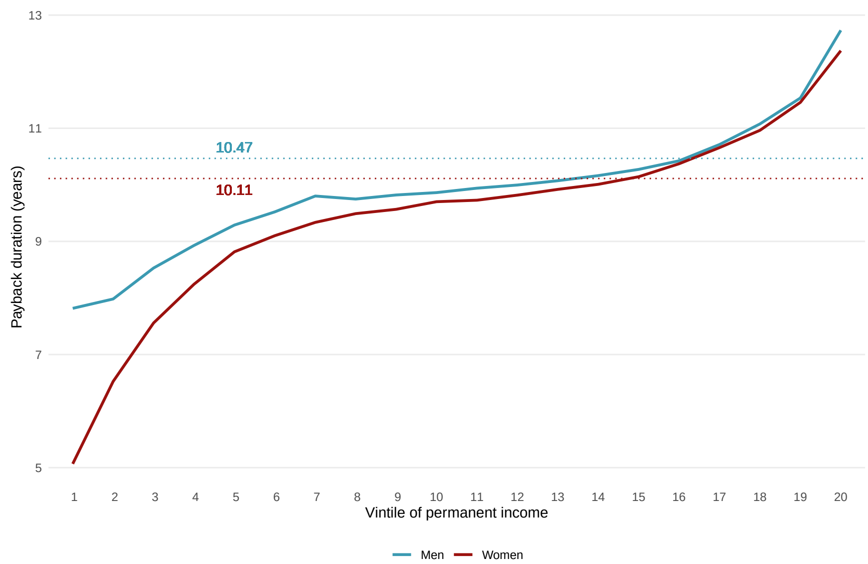
[20]Relative pension levels above 100% might seem counterintuitive. It suggests that differential in contributions between men and women is more than compensated by the pension formula, suggesting no contributivity at all. It might be due to some familial rights that affect mostly women.

Figure F.1: Contribution and benefit of women relative to men, by income group



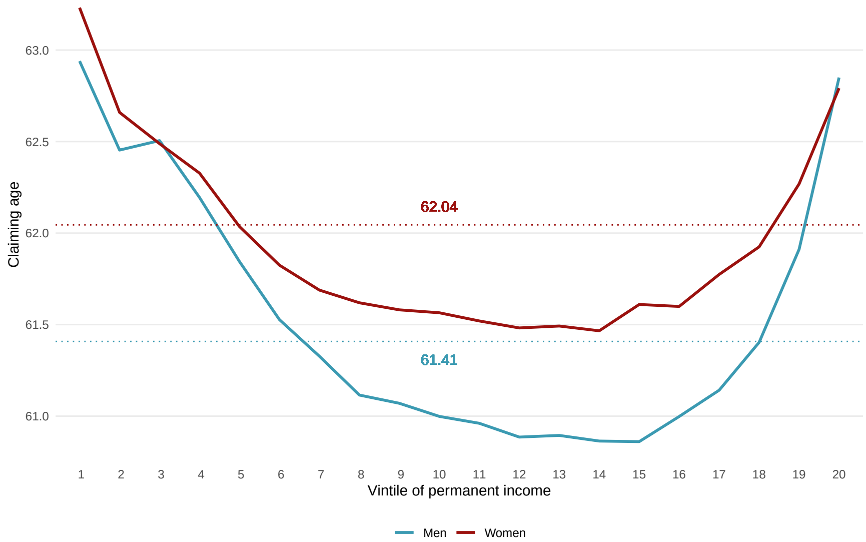
Note : This figure presents average pension level and amount of lifetime contributions of women by vintile of labor permanent income, relatively to the same measure among men of the same income group. Income rank is defined for the whole sample – men and women pooled. Discount factor used in computation of lifetime contributions equals 1.
Interpretation: At the 3^{rd} vintile, lifetime pension contribution of women are around 20% lower than men's on average, while pension levels are equal on average.
Source : Agirc-Arrco data, authors' calculations

Figure F.2: Payback duration by gender and income group



Note : This figure presents average payback period by vintile of labor 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 13^{th} vintile need to receive pensions during 10 years for lifetime pensions to equal lifetime contributions – $RC = 100\%$.
Source : Agirc-Arrco data, authors' calculations

Figure F.3: Average claiming age by gender and income group



NOTE : This figure presents average observed claiming age by vintile of labor 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} vintile, average claiming age of men is 0.6 years lower than retirement age of women.
SOURCE : Agirc-Arrco data, authors' calculations

G Implicit tax/subsidy rates

In its more general form, we defined RC as :

$$RC_x = \frac{\sum_{a=55}^{120} \beta^a P_{xa}}{\sum_{a=8}^{68} \beta^a C_{xa}}$$

where C_{xa} and P_{xa} are respectively total contributions and total pensions of individuals of group x at age a. We quantify the implicit redistribution generated by mortality differentials. To do so, we consider a benchmark scenario in which all individuals of x follow the same mortality pattern as the average population of reference – all individuals of a given gender, or the overall population, depending on the analysis. In this benchmark scenario, return on contributions for x is different and denoted $\widetilde{RC}_x$. We then define the implicit tax rate on contributions τ_x as the tax rate on observed contributions that rationalises the observed flow of pensions, under the return $\widetilde{RC}_x$:

$$\widetilde{RC}_x = \frac{\sum_{a=55}^{120} \beta^a P_{xa}}{\sum_{a=8}^{68} \beta^a C_{xa}(1 - \tau_x)}$$

One also has :

$$\tau_x = 1 - \frac{RC_x}{\widetilde{RC}_x}$$

One can interpret a postive tax rate τ_x as the share of pensions that virtually gives access to no pension right, compared to the benchmark return on contributions, and because of lower longevity than average. Reversely, a negative tax rate is an implicit subsidy on contributions through higher longevity than average. Finally, note that a τ_x tax on contributions can also be interpreted as a τ_x implicit tax on pensions.