Decomposing Changes in Income Mobility Over Time

FRANCISCO LIBANO-MONTEIRO*

London School of Economics

Abstract

An economy's income mobility can change substantially over time. However, a substantial decline in population income mobility may either be the result of a decline in mobility within population subgroups, or of a growing share of population subgroups that are less mobile, such as women and older workers. This paper offers an exact decomposition of changes in mobility over time. Population income mobility is not a simple weighted average of subgroup mobilities, and can even be lower than that of every subgroup, a mild form of Simpson's paradox. As a result, beyond compositional dynamics in the labor force, the dynamics of inequality and of between-group gaps also matter for aggregate mobility. I apply the decomposition to Portugal using administrative data on the universe of private-sector workers. Between 1986 and 2019, population wage mobility fell by around 23 percent, yet it remained broadly stable within age-gender-education groups, revealing the importance of structural differences between groups, such as their labor-force shares and income gaps. Finally, I find a Great Gatsby relationship across subgroups, with more unequal groups being less mobile, gender being the exception. Across time, however, inequality and intragenerational mobility are essentially unrelated.

Keywords: Income Mobility, Subgroup Decomposition; Wage Dynamics; Structural Change
JEL codes: D31; J31; J62

*Contact: M.F.LIBANO-MONTEIRO@LSE.AC.UK. I am deeply grateful to my supervisors, Wouter Den Haan, Xavier Jaravel, and Ricardo Reis, for their guidance and support. I also thank Francis Kramarz, Frank Cowell, and Chad Syverson for early discussions. I thank participants at the LSE Macroeconomics work-in-progress seminar, the NHH Bergen workshop, and the INTEDU Unimore Modena conference for helpful comments. Phyu Sin Moe provided excellent research assistance. I gratefully acknowledge financial support from the Fundação para a Ciência e a Tecnologia (FCT). I thank GEP and DGEEC for facilitating access to the Quadros de Pessoal data.

1 Introduction

Research on (relative) income mobility, the extent to which people's positions in the distribution churn, is a central object in the study of inequality. While income mobility has long been studied with richer objects such as transition matrices or the income copula, it is commonly summarized in a single scalar, such as the rank-rank slope or the correlation coefficient. Single-scalar measures are widely used in both the intergenerational income mobility literature, which compares a child's income rank with their parents', and the intragenerational one, which compares a worker's income rank today with their rank years later [Jäntti and Jenkins, 2015, Guvenen et al., 2022]. A single number allows easy comparison across countries and over time. As administrative data accumulate, it becomes increasingly possible to ask not only how mobile a population is, but how its income mobility changes over time.

Population income mobility, however, hides substantial heterogeneity across population subgroups. In the intergenerational context, children of poor parents in some U.S. commuting zones have a real chance of reaching the top of the distribution while those in others overwhelmingly remain poor. A child with parents at the 25th percentile reaches an expected rank of 46 in Salt Lake City, the most upwardly mobile of the fifty largest commuting zones, and only 36 in Charlotte, the least [Chetty et al., 2014]. In the intragenerational context, the same holds across groups within a labor market. In the Portuguese data studied below, low-wage women are less likely to climb the wage distribution over the following five years than low-wage men. A single population number averages over all of this and, as this paper shows, does so in a way that is far from transparent.

This matters because a change in the population number can have two very different sources. Income mobility may have fallen within subgroups, or population weight may simply have shifted toward subgroups that were already less mobile. Portugal makes the ambiguity concrete. Between 1986 and 2019 its wage mobility fell substantially, while its labor force was transformed: more women entered employment, the workforce aged, schooling rose several-fold, and jobs moved from goods into services [Oliveira, 2024]. Since women and older work-

ers have more persistent wages, a natural hypothesis is that the decline is a compositional artifact. The population time series alone cannot tell these apart.

This paper separates the forces behind changes in population income mobility into a small number of channels, and takes them to the universe of Portuguese private-sector workers. As Section 3.1 shows, the level of population income mobility is not a simple weighted average of subgroup income mobilities [Hertz, 2008, Jácome et al., 2025]. As a result, beyond compositional dynamics, the dynamics of other structural forces, such as the gaps between groups, also matter for changes in population income mobility. The intuition is the following. If the gender gap widens substantially, then even if men and women churn among themselves just as much as before, there is less churning across the two groups, making overall population income mobility smaller. The paper's main contribution is a closed-form identity that splits the change in population income mobility over any period into four channels: within-group income mobility, composition, inequality, and the persistence of between-group gaps. The four sum to the total change exactly.

While this decomposition applies to any slope-based income mobility measure, and indeed to any continuous outcome variable, in this paper I look at changes in intragenerational wage mobility in Portugal, where the wage is income from labor (earnings) divided by hours.[1] I use *Quadros de Pessoal*, an administrative matched employer–employee census that covers every private-sector firm and carries the worker characteristics (gender, age, education) and firm characteristics (region, sector) needed to partition the population in many ways.[2] Population income mobility moved substantially over the period. The five-year rank-rank slope rises from 0.720 in 1986 to 0.810 around 2010–2012 before receding to 0.785 by 2014.

Portugal makes an informative case. By the 2000s it ranked among the least mobile economies in Europe, on household income as on men's wages [Sologon and O'Donoghue, 2009, 2012, Van Kerm and Pi Alperin, 2013, Alves and Martins, 2012, OECD, 2018], whereas

[1] The same decomposition applies to labor income mobility, which combines wage mobility and hours mobility.

[2] This panel contrasts with other major administrative data sources, such as those in the UK and the US, which also have high-quality earnings data but limited information on hours worked.

in the late 1980s its wage mobility was similar to the United Kingdom's [Cardoso, 2006]; on the measure of the Global Repository of Income Dynamics, computed on the Portuguese data, it moved from the level of the most mobile economies in the late 1980s to the lower third of the comparison by 2007 (Guvenen et al., 2022; Appendix H). Over 1986 to 2019 its five-year wage persistence rose and then only partly receded, leaving population income mobility around 23 percent lower at the end of the period than at the start, or equivalently wage persistence around 10 percent higher.[3]

The main finding is that income mobility within finely defined groups barely moved. In particular, grouping workers jointly by age and education suffices to account for the change in population income mobility. Within these groups, income mobility is essentially flat, so the decline is driven by structural shifts between them. Looking at any single grouping alone, however, by gender, age, education, region, or sector separately, cannot account for the decline.

Lastly, I investigate two Great Gatsby relationships. The first, an intra-group Great Gatsby curve, asks whether groups with lower income mobility are also groups with higher income inequality; the second, an intertemporal Great Gatsby curve, asks whether periods of lower income mobility are also periods of higher income inequality. The first holds, more unequal groups being markedly less mobile, with gender the exception; the second does not, since over time wage dispersion and wage persistence are essentially unrelated in Portugal.

Related literature. This paper builds on a decomposition of the *level* of income mobility due to Hertz [2008] and applies it empirically across population subgroups. Its main contribution is to take the further step of decomposing *changes* in income mobility over time and applying it to the evolution of intragenerational mobility in Portugal.

Hertz [2008] decomposes a population regression slope into a weighted subgroup component and a between-group gap term. I apply this decomposition to the Portuguese data to uncover the heterogeneity in income mobility across a variety of subgroups and to gauge the

[3]For the rank-rank measure, mobility is $1-\rho$. The rise in persistence from 0.720 to 0.785 reduces mobility from 0.280 to 0.215. The same change of 0.065 is about 9% of initial persistence and 23% of initial mobility.

contributions of the subgroup component and the group-gap term.

A separate, axiomatic literature criticizes the statistics that Hertz-type decompositions take as given. Cowell and Flachaire [2018, 2025] characterize indices of how far individual status moves between two dates and, from that standpoint, judge the elasticity and the correlation to be neither measures of movement nor decomposable in the additive sense of Shorrocks [1980]. The object here is positional persistence rather than distance travelled [McGee, 2026].

A large literature documents how differently income mobility runs across population subgroups. In the *intergenerational* context, this heterogeneity is well established, across space and race [Chetty et al., 2014, 2020]. In the *intragenerational* context it is studied far less; Guvenen et al. [2022] report income mobility separately for men and women, and for younger and older workers, across countries, and Buchinsky et al. [2003] compare earnings mobility across subgroups by gender and education in France. Applying the decomposition to Portugal, I find that the gap term makes an important contribution to population persistence, particularly where income gaps between groups are large, generating a mild form of Simpson's paradox.[4] The paper's first contribution is thus a systematic, exact subgroup decomposition of intragenerational income mobility, applied to administrative data on the universe of Portuguese private-sector workers.

The second contribution is an exact decomposition of *changes* in income mobility over time, which lets me ask whether structural shifts across these groups can account for changes in population income mobility, attributing the change to a small number of channels. To my knowledge no prior work decomposes changes in intragenerational income mobility in this way. Guvenen et al. [2022] document the level and trend of intragenerational mobility across countries, and Buchinsky et al. [2003] track earnings-mobility trends over time for France as a whole and for subgroups, but neither decomposes the change. The closest decomposition exercise is intergenerational. Jácome et al. [2025] apply the Hertz level identity to rank-rank

[4] Across education groups, population mobility is lower than mobility within each group taken separately, echoing Torche and Corvalán [2018], who show that a persistence coefficient estimated on group means, surnames in their case, is the between-group component of the individual coefficient and loads persistence onto the differences between group means.

mobility and attribute the rise in U.S. intergenerational mobility between the 1910s and 1940s cohorts through counterfactual simulations that switch off one force at a time, for instance holding the white-only slope or the between-group term at its base-period value, rather than in an exact, closed-form decomposition that sums to the total change. The identity derived here instead attributes the entire change to four channels that sum to it exactly. Karlson and In [2025] decomposes changes in sibling correlations, a related but distinct exercise for a different measure; Lee and Solon [2006] study trends in intergenerational elasticities without a subgroup decomposition.

On the Portuguese labor market itself, Cardoso [2006] documents the decline in wage mobility; mine is the first paper to trace its anatomy at this level of demographic granularity and to decompose what accounts for that decline. Relatedly, Carneiro et al. [2023] decompose the *level* of wage persistence in these data into worker, firm, and match components, finding that unobserved worker heterogeneity accounts for most of it; I instead ask whether changes in composition and group structure account for the change in aggregate persistence over time.

Lastly, I relate my findings to the Great Gatsby curve. Where that relationship is usually drawn across countries and across generations, linking higher income inequality to lower income mobility [Corak, 2013, Guvenen et al., 2022], I draw it within a single economy: an intra-group Great Gatsby curve across subgroups, and an intertemporal Great Gatsby curve across periods. The first holds in the Portuguese data and the second does not, a contrast that a permanent–transitory view of income makes natural (Appendix H.1).

Outline. Section 2 introduces the mobility measures, describes the data, and defines the subgroup partitions. Section 3 derives the level decomposition, provides intuition for the resulting formula, and derives the changes decomposition. Section 4 applies the framework to the Portuguese mobility decline. Section 5 discusses the interpretation of the decomposition. Section 6 concludes.

2 Income Mobility Measures, Data and Subgroup Partitions

2.1 Income mobility measures

Individuals occupy different positions in the income distribution over their lifetimes. The degree to which they change position is what the literature calls *relative* income mobility, and it can be summarised in several ways. The copula of the joint distribution of $(y_{i,t}, y_{i,t+h})$ retains the full dependence structure; transition matrices coarsen it into movements between quantile bins; and scalar measures condense it into a single number [Cowell and Flachaire, 2018, Jäntti and Jenkins, 2015]. Each choice trades one thing for another. The copula and the transition matrix preserve information about *where* in the distribution churning happens, whether the bottom of the distribution is more or less fluid than the top, but they are unwieldy to compare across populations or to track over time. A scalar measure gives up that detail. It reports an aggregate degree of churning without distinguishing intensity at different points of the distribution. What it buys is comparability, which is why scalar measures, the rank-rank slope [Chetty et al., 2014], the intergenerational elasticity [Lee and Solon, 2006], the correlation coefficient, are the objects whose time series are reported, compared across countries, and debated. This paper offers a decomposition of these scalar measures.

Throughout, $y_{i,t}$ denotes an income measure for individual i at date t, a percentile rank, a log wage, or any real-valued outcome, and mobility is measured at horizon h by relating $y_{i,t+h}$ to $y_{i,t}$ among the individuals observed at both dates.[5] The two workhorse measures are the *slope* of the regression of $y_{i,t+h}$ on $y_{i,t}$,

$$\rho_{t,h} = \frac{\text{Cov}(y_{i,t+h},\, y_{i,t})}{\text{Var}(y_{i,t})} = \frac{\text{Cov}(y_{i,t+h},\, y_{i,t})}{\sigma_t\, \sigma_t}, \tag{1}$$

[5] In the empirical application I use a five-year horizon ($h = 5$) as the baseline, following the GRID convention [Guvenen et al., 2022]. I also check robustness using a three-year horizon ($h = 3$), and the results are qualitatively similar.

and the *correlation coefficient*,

$$r_{t,h} = \frac{\text{Cov}(y_{i,t+h},\, y_{i,t})}{\sigma_t\, \sigma_{t+h}} = \rho_{t,h} \times \frac{\sigma_t}{\sigma_{t+h}}, \tag{2}$$

where $\sigma_t^2 = \text{Var}(y_{i,t})$. Written this way, the two measures are visually one substitution apart. The slope divides the covariance by $\sigma_t \sigma_t$, the correlation by $\sigma_t \sigma_{t+h}$. When $y_{i,t}$ is a percentile rank, (1) is the rank-rank slope; when $y_{i,t}$ is a log wage, it is the intragenerational elasticity. Both are *immobility* measures. Higher values mean more persistence and less churning. The same objects applied to parent–child pairs measure intergenerational mobility; everything below applies unchanged.

The choice between the two measures interacts with the choice of income concept. The natural requirement is a measure that equals one when relative positions are unchanged. The intragenerational elasticity (or log-log slope) does not deliver this. If dispersion rises between t and $t+h$ while everyone keeps their relative position, the slope will not be one. This is why, for log wages, the natural measure is the correlation coefficient[6], which normalizes by $\sigma_t \sigma_{t+h}$ and so stays at one under unchanged positions. Appendix C works through a simple AR(1) example in which a pure rise in churning leaves the slope unchanged but lowers the correlation, isolating the role of dispersion. For percentile ranks the issue does not arise. Ranks fix the dispersion at every date[7], so σ_t / σ_{t+h} is one and the slope already coincides with the correlation.

[6] Equivalently, the correlation coefficient is the slope obtained after standardizing income at each date. Regressing the z-score of $y_{i,t+h}$ on the z-score of $y_{i,t}$ returns r directly, since standardization resets the cross-sectional dispersion to one at both dates. This is why the correlation, unlike the slope, is stable to shifts in dispersion between t and $t+h$, and why decomposing changes in the correlation isolates the contribution of shifting dispersion as a distinct term.

[7] This holds exactly when incomes are distinct. When many workers share the same income, as at the minimum wage, they are assigned a common rank, so the rank distribution has mass points rather than being exactly uniform, and the slope and correlation coincide only approximately. In the data, between 0.6% and 3.1% of workers share the modal wage in a given year, and assigning tied workers their average rank, the lowest rank, the highest rank, or a random order changes the rank-rank slope by at most 0.001 in any base year.

2.2 Data

I use data from the Portuguese *Quadros de Pessoal* (QP), an annual matched employer–employee census conducted by the Ministry of Employment. Every firm with at least one salaried worker is legally required to report its full roster of employees each October, along with detailed information on worker compensation, hours, demographics, and firm characteristics. To reduce misreporting, firms must also post their employee roster, including names, job titles and monthly pay, inside their premises. The Ministry distributes an electronic version of the data with longitudinal identifiers for each worker, establishment and firm, enabling matched employer–employee analysis.

QP covers the universe of private-sector employment in Portugal. It excludes public-sector employees (approximately 15% of total employment) and the self-employed (approximately 18%), capturing around two-thirds of total employment. In 2006, for example, the raw data contain 3.1 million workers against 4.75 million employed individuals in the Labour Force Survey. Longitudinal worker identifiers allow individual workers to be tracked across firms and over time.

The analysis window is 1986–2019. The annual coverage has two gaps. First, the QP survey was not conducted in 1990 or in 2001, so those two years contain no observations at all. Second, workers' completed schooling is not recorded in 1995, 1997, or 1998, so any decomposition that classifies workers by education omits those three years. I work with a cleaned 10% worker-level random sample, approximately 6.0 million observations, of which 5.5 million carry valid wages, covering around 730,000 unique workers. In figures, I linearly interpolate across the missing years; in bar charts the interpolated years are drawn lighter.

The data contain rich information at the worker level (base earnings, overtime pay, regular and irregular bonuses, normal and overtime hours, age, gender, education, four-digit occupation and job title, tenure, and part-time status) as well as at the firm level (region, four-digit industry, firm size, and annual revenue). For the analysis I measure total compensation as the sum of base earnings and all bonus components, and construct real hourly earnings, wages, by

dividing total compensation by total hours, normal plus overtime, deflated to 2019 euros using the CPI. I exclude observations with zero or missing reported hours. The outcome variables are the log real wage and percentile ranks. Ranks are a transformation of a given vector of data indexed from 0 to 1, where 0 represents the lowest-earning individual and 1 the highest-earning. If rank repetition does not exist, the distribution is uniform. I also report results for "permanent" income, defined as the centred three-year mean of the log real wage [following Guvenen et al., 2022], which smooths transitory fluctuations.

This richness at both the worker and the firm level matters for what follows. It is what allows the workforce to be partitioned along many dimensions, gender, age, education, geographical region, and broad industry, which is precisely what the decomposition framework of Section 3 requires. The construction of each partition is detailed in Section 2.3. Beyond richness, the administrative nature of the data is of independent importance for the measurement of mobility. Key for my analysis is the covariance between current and future income, and measurement error would shrink this covariance and thus decrease the measured persistence. More perniciously for the question studied here, any *improvement* in data quality over time would produce a spurious rise in measured persistence even if the underlying wage dynamics were unchanged. QP is largely immune to this concern. Earnings are drawn from firms' payroll records under a legal reporting obligation, and firms must publicly post the reported rosters, discouraging misreporting throughout the sample period. I therefore interpret the rise in measured persistence documented below as a genuine change in wage dynamics rather than an artifact of declining measurement error. The permanent-income specification, which averages log wages over three adjacent years and thereby purges transitory fluctuations and residual noise, provides a further check. The decline in mobility is, if anything, more pronounced for permanent income (Figure 16).

2.3 Subgroup partitions

The decomposition of Section 3 requires the workforce to be split into mutually exclusive groups. I use five one-dimensional partitions and their joint partitions. In every case a worker is assigned to a group by their characteristics in the *base* year t, so group membership never conditions on later outcomes (Section 3 discusses this choice).

The five one-dimensional partitions are as follows. *Gender* splits the workforce into men and women, as recorded. *Age* uses equal-share terciles of base-year age, "young" (≤ 31), "middle" (32–43), and "old" (≥ 44), with the cutoffs chosen so that each tercile holds around one-third of the pooled sample. *Education* groups workers by completed schooling into three levels: basic (≤ 6 years), secondary (7–12 years), and higher (> 12 years). *Region* collapses the NUTS-2 classification into three broad areas, Lisbon, the North, and a residual Centre–South that pools the remaining mainland and island regions. *Industry* splits the one-digit industry classification into two broad sectors, Goods (sections A–F: agriculture, mining, manufacturing, utilities, and construction) and Services (sections G–U). Workers whose region or industry is missing or unclassifiable are dropped from the corresponding partition.

To capture interactions across dimensions, I also form joint partitions: all ten pairs of the five characteristics, and the full gender×age×education cut. The partitions that carry the main results combine age and education, the two characteristics with the largest wage gaps between groups. In the three-way cut, age enters as the same three terciles used elsewhere, so the population divides into $K = 18$ cells, {young, middle, old}×{women, men}×{basic, secondary, higher}, which average around 5,700 workers per base year, although the smallest cells of the 1980s hold only a few dozen workers (Appendix E, Table 5); the nine-cell age×education partition, whose smallest cell holds 315 workers, carries the paper's main result. The partitions that use education inherit the education-complete flow, so they cover only base years in which schooling is recorded.

3 Theory: Subgroup Decompositions

I start with the decomposition of the *level* of population mobility, provide intuition for the resulting formula, and then extend it to a decomposition of *changes* in mobility over time, the tool that Section 4 applies to the large changes in intragenerational mobility in the Portuguese labor market.

3.1 Decomposing the level of income mobility

Partition the population into two mutually exclusive groups, with size $N_M + N_W = N$ and population shares $n_g = N_g/N$; I use gender, men M and women W, as a running example for clarity, but everything that follows applies to any partition of the population. The decomposition also places no restriction on *how* the partition is constructed. Groups may be defined by any function of worker characteristics, observed at t, at $t+h$, or both. The partition may equally be cross-dimensional, its cells formed by interacting several characteristics at once, for instance age, gender, and education jointly; Section 4.2 and Appendix E.2 report such an age×gender×education decomposition. In the empirical work, I classify workers by their characteristics at the base year t.[8] For a group g, write the group means as $\bar{y}_{g,s}$, the group variances as $\sigma^2_{g,s}$ for $s \in \{t, t+h\}$, and the group covariance as $\text{Cov}_g = \text{Cov}_g(y_{i,t+h}, y_{i,t})$; population moments carry no group subscript. Define the mean gap between the two groups at each date,

$$\bar{y}_s^{GAP} = \bar{y}_{M,s} - \bar{y}_{W,s}, \qquad s \in \{t,\ t+h\}.$$

[8] For example, workers can be classified as having basic, secondary, or higher education at time t, regardless of their education at $t+h$; alternatively, the partition could interact the two dates, basic education at both dates, basic at t and secondary at $t+h$, secondary at t and higher at $t+h$, and so on. Partitions may also be joint in several characteristics, young men, young women, older men, and older women, for example. I adopt base-year classification for two reasons. First, group membership is then not conditioned on subsequent outcomes, so each group's mobility retains its forward-looking interpretation. How do workers who looked identical at t fare over the following h years? Second, while a partition interacting the two dates preserves every worker, switchers simply populate their own cells, to the extent that switching is rare, many switcher cells would contain few observations. More importantly, within-group churning in such cells is harder to interpret. One can readily imagine workers of the same education background or the same gender holding and trading similar jobs, whether through skills or preferences, but it is much harder to attach meaning to churning among workers who have basic education at t and secondary education at $t+h$.

The general case with K groups is presented in Appendix A; every object below generalizes with the single gap term becoming a sum over group pairs.

Derivation. Write the covariance in (1) as a population average and clear the $1/N$:

$$\rho_{t,h} = \frac{\sum_i (y_{i,t+h} - \bar{y}_{t+h})(y_{i,t} - \bar{y}_t)}{N \ \mathrm{Var}(y_{i,t})}. \tag{3}$$

Split the numerator over $i \in M$ and $i \in W$, and inside each factor add and subtract the group's own mean. For women,

$$\sum_{i \in W} (y_{i,t+h} - \bar{y}_{t+h})(y_{i,t} - \bar{y}_t) = \sum_{i \in W} \big[(y_{i,t+h} - \bar{y}_{W,t+h}) + (\bar{y}_{W,t+h} - \bar{y}_{t+h})\big]\big[(y_{i,t} - \bar{y}_{W,t}) + (\bar{y}_{W,t} - \bar{y}_t)\big].$$

Expanding gives four sums, of which the two cross terms vanish because deviations from a group's own mean sum to zero within the group. What survives is a subgroup covariance piece and a mean-gap piece:

$$\sum_{i \in W} (y_{i,t+h} - \bar{y}_{t+h})(y_{i,t} - \bar{y}_t) = N_W \, \mathrm{Cov}_W + N_W (\bar{y}_{W,t+h} - \bar{y}_{t+h})(\bar{y}_{W,t} - \bar{y}_t),$$

and analogously for men. Because the pooled mean is the share-weighted group mean, $\bar{y}_s = n_M \bar{y}_{M,s} + n_W \bar{y}_{W,s}$, each group's deviation from the pooled mean is proportional to the gap between the groups:

$$\bar{y}_{W,s} - \bar{y}_s = -n_M \, \bar{y}_s^{GAP}, \qquad \bar{y}_{M,s} - \bar{y}_s = n_W \, \bar{y}_s^{GAP}.$$

Substituting and collecting, the two mean-gap pieces combine into a single gap term:

$$\sum_i (y_{i,t+h} - \bar{y}_{t+h})(y_{i,t} - \bar{y}_t) = N_M \, \mathrm{Cov}_M + N_W \, \mathrm{Cov}_W + \frac{N_M N_W}{N} \, \bar{y}_{t+h}^{GAP} \, \bar{y}_t^{GAP}. \tag{4}$$

The denominator is the same object with $y_{i,t+h}$ replaced by $y_{i,t}$, a variance is the covariance of a variable with itself, so

$$N \text{ Var}(y_{i,t}) = N_M \sigma^2_{M,t} + N_W \sigma^2_{W,t} + \frac{N_M N_W}{N} \left(\bar{y}^{GAP}_t\right)^2. \tag{5}$$

Finally, divide (4) by (5) term by term, using the identity $\frac{A+B+C}{D+E+F} = \frac{A}{D}\frac{D}{D+E+F} + \frac{B}{E}\frac{E}{D+E+F} + \frac{C}{F}\frac{F}{D+E+F}$. Each numerator piece, divided by its matching denominator piece, is a *rate*, and each denominator piece, divided by the total, is a *weight*.

Proposition 1 (Level decomposition of the slope, $K = 2$)**.**

$$\boxed{\rho_{t,h} = \omega_M \rho_M + \omega_W \rho_W + \omega^{GAP} \frac{\bar{y}^{GAP}_{t+h}}{\bar{y}^{GAP}_t}} \tag{6}$$

where $\rho_g = \text{Cov}_g / \sigma^2_{g,t}$ *is the within-group slope for group* g *and*

$$\omega_g = n_g \frac{\sigma^2_{g,t}}{\sigma^2_t}, \qquad \omega^{GAP} = n_M n_W \frac{\left(\bar{y}^{GAP}_t\right)^2}{\sigma^2_t}, \qquad \omega_M + \omega_W + \omega^{GAP} = 1.$$

Population immobility is thus the sum of three terms. The first two, each group's within-group persistence weighted by ω_g, form the *weighted subgroup component*, or simply the *subgroup component*. The third is a gap term, the ratio of the end-period to the base-period mean gap, weighted by ω^{GAP}. The subgroup component depends on both the within-group coefficients and their population and variance weights. The weights sum to one exactly, by (5). The gap weight is where the gap enters. If there is no gender gap at the base date ($\bar{y}^{GAP}_t = 0$), then $\omega^{GAP} = 0$ and population immobility reduces to a genuine weighted average of the subgroup slopes.[9] A larger gap means a larger ω^{GAP}, and mechanically less scope for churning across subgroups. If men and women occupied increasingly separate segments of the income distri-

[9] When the base-year gap is zero the ratio $\bar{y}^{GAP}_{t+h} / \bar{y}^{GAP}_t$ is undefined, but the gap term itself, $\omega^{GAP}\, \bar{y}^{GAP}_{t+h} / \bar{y}^{GAP}_t = n_M n_W\, \bar{y}^{GAP}_t\, \bar{y}^{GAP}_{t+h} / \sigma^2_t$, is well defined and equals zero; the decomposition is always evaluated in this product form. With many groups, a near-zero base-year gap between one pair makes that pair's ratio large and its weight tiny, while their product, and the gap term as a whole, are unaffected.

bution, reshuffling would happen only within each subgroup while positions across segments stayed frozen, raising population immobility.

The correlation coefficient. The correlation (2) has the same numerator (4); only the normalization changes, from $\sigma_t\sigma_t$ to $\sigma_t\sigma_{t+h}$. Because the cross-period product $\sigma_t\sigma_{t+h}$, unlike a single variance, does not split into group pieces, the three numerator terms are divided by it directly, inserting each group's own dispersion to form rates and weights. The result has exactly the same architecture:

$$\boxed{r_{t,h} = \omega^c_M\, r_M + \omega^c_W\, r_W + \omega^{c,GAP}\,\frac{\bar{y}^{GAP}_{t+h}}{\bar{y}^{GAP}_{t}}} \tag{7}$$

with $r_g = \mathrm{Cov}_g\,/(\sigma_{g,t}\sigma_{g,t+h})$ the within-group correlation and

$$\omega^c_g = n_g\,\frac{\sigma_{g,t}\,\sigma_{g,t+h}}{\sigma_t\,\sigma_{t+h}}, \qquad \omega^{c,GAP} = n_M n_W\,\frac{\left(\bar{y}^{GAP}_t\right)^2}{\sigma_t\,\sigma_{t+h}}.$$

The first two terms form the correlation counterpart of the subgroup component. The only substantive difference is that the variance share $\sigma^2_{g,t}/\sigma^2_t$ becomes a *dispersion share* $\sigma_{g,t}\sigma_{g,t+h}/(\sigma_t\sigma_{t+h})$, and, because the denominator no longer splits additively, the correlation weights do not sum to one. The full derivation, and the check that (7) equals (6) times σ_t/σ_{t+h}, are in Appendix B.

3.1.1 The level decomposition: interpretation and heterogeneity

Two properties of the decomposition carry its economic content.

Observation 1: population mobility is not a weighted average of subgroup mobilities. The population slope ρ is not a weighted average of the subgroup slopes ρ_M and ρ_W, and so need not lie between them. There is a third term, the persistence of the mean gap between the groups, $\bar{y}^{GAP}_{t+h}/\bar{y}^{GAP}_t$, carrying weight ω^{GAP}. To see why it matters, suppose men and women

churned to exactly the same extent, so that $\rho_M = \rho_W$. Even then, if the two groups earn very different amounts on average, few women climb into the male part of the distribution and few men fall below the women's, so there is little income mobility *across* groups, even when there is a great deal of it *within* each group. Positions stay persistent across groups no matter how fluid they are within them. Churning is of two kinds: workers reshuffling within their group, and workers crossing between groups. The first two terms capture the within-group kind, the third the across-group kind.[10] The wider the gaps between groups, the less workers cross between them, and the less income mobility there is overall.

A mild Simpson's paradox. That the pooled slope can differ sharply from every within-group slope is a familiar phenomenon. The slope relating a variable Y to a variable X can change, or even reverse, once one conditions on a third variable Z. Here Y is the individual worker's income rank $y_{i,t+h}$, X is the rank $y_{i,t}$, and Z is the subgroup. So the population slope of $y_{i,t+h}$ on $y_{i,t}$ need not resemble the slopes estimated within groups. This is an instance of Simpson's paradox.[11] My data display the milder, magnitude version (Table 1). For gender the paradox does not arise, as the empirics below confirm. The population slope sits between the two subgroup slopes, 0.785 in 2014 against 0.736 for men and 0.810 for women, though the gender gap term is far from negligible, accounting for 0.083 of the population slope in 1986 and 0.040 by 2014. For education it does arise, in mild form. The population slope lies above *every* education group's slope, 0.785 against 0.740 for basic, 0.710 for secondary, and 0.703 for higher education, because the large and persistent education wage gap, the skill premium, keeps workers from crossing between education groups even as they churn within their education cells. This foreshadows the decomposition of changes. Within-group mobility may stay constant over time, yet aggregate mobility can still move if between-group gaps shift,

[10] To track how average within-group mobility evolves over time, one can first residualize income on a full set of base-year subgroup indicators in the paired sample and compute the persistence slope in residualized income. This average uses population and variance weights that may change over time.

[11] In its classic form the slope sign reverses. Every within-group slope is negative, yet the pooled slope comes out positive, because pooling adds in the gap term. Appendix D walks through both this classic sign-reversal case and the milder configuration relevant here, in which the pooled slope keeps its sign but lies outside the range of the group slopes.

for instance a widening education wage premium.

Observation 2: the weights are composition-times-inequality weights, not composition weights. The subgroup weight $\omega_g = n_g \sigma^2_{g,t}/\sigma^2_t$ is the population share multiplied by the group's share of income variance. A group's imprint on population mobility is therefore governed not only by how many workers it contains, but by how much of the income distribution it spans. Suppose men's incomes are spread across the whole distribution while women's are bunched tightly together. Then women occupy a single narrow slice, few of them ever cross into the men's range, and even if they churn heavily among themselves, the single scalar mobility measure barely moves, though a transition matrix would register their churning.[12] The aggregate slope is governed by the men, who span the distribution, almost regardless of how many women there are.[13] A small but internally unequal group, older workers with heterogeneous careers, say, can carry weight far beyond its headcount, while a large group compressed into a narrow income band carries less than its share. This distinction is exactly what the *inequality* channel of the changes decomposition below isolates, and it matters empirically. In the education partition, the basic-education group's weight collapses from 0.72 to 0.20 over the sample not only because the group shrank, but because its variance share fell as the educated groups came to own most of the wage dispersion.

3.2 Decomposing changes in income mobility

Suppose one is interested not only in how much individuals churn through the income distribution from a base year t to $t+h$, but in how much the level of that churning changes over time and in what can account for those changes. The level decomposition writes population

[12] With only men, the dots lie along the 45° line, slope one; now add the women as a tight cloud in one spot, which reshuffles internally from period to period but stays put, so the fitted line hardly tilts and the slope stays near one.

[13] This logic is about dispersion in the units of the regression, and applies to the slope of log income. Under the rank-rank slope the picture differs. Ranks are uniform by construction, so a large group clustered in income still spreads across much of the rank distribution, and therefore carries a large share of the rank variance and a large weight.

mobility at any date as a sum of weight-times-rate products,

$$\rho = \sum_a W_a R_a, \qquad a \in \{M,\ W,\ GAP\},$$

with W_a the weights and R_a the rates (the within-group slopes and the gap ratio) of Proposition 1. Comparing two base years t_0 and t_1 at a common horizon h, and writing $\Delta x = x_{t_1} - x_{t_0}$ with subscripts t_0, t_1 for quantities evaluated at each date, the change in population mobility is a telescoping sum of products. The engine that splits it is the exact *product–difference identity*

$$W_{t_1} R_{t_1} - W_{t_0} R_{t_0} = W_{t_0}\,\Delta R + \Delta W\, R_{t_1}, \tag{8}$$

the base-year weight on the rate change plus the end-year rate on the weight change. Applying (8) to each of the three terms of (6), and further splitting each subgroup weight change, since $\omega_g = n_g v_g$ is itself a product of the population share n_g and the variance share $v_g = \sigma^2_{g,t}/\sigma^2_t$, so that $\Delta\omega_g = \Delta n_g\, v_{g,t_1} + \Delta v_g\, n_{g,t_0}$, yields the paper's central identity.

Proposition 2 (Changes decomposition of the slope, $K = 2$)**.**

$$\begin{aligned}\Delta\rho = {} & \underbrace{\omega_{M,t_0}\,\Delta\rho_M + \omega_{W,t_0}\,\Delta\rho_W}_{\text{within}} + \underbrace{\Delta n_M\, v_{M,t_1}\,\rho_{M,t_1} + \Delta n_W\, v_{W,t_1}\,\rho_{W,t_1}}_{\text{compositional}} \\ & + \underbrace{\Delta v_M\, n_{M,t_0}\,\rho_{M,t_1} + \Delta v_W\, n_{W,t_0}\,\rho_{W,t_1}}_{\text{inequality}} + \underbrace{\omega^{GAP}_{t_0}\,\Delta r^{GAP} + \Delta\omega^{GAP}\, r^{GAP}_{t_1}}_{\text{group gap}},\end{aligned} \tag{9}$$

where $r^{GAP} = \bar{y}^{GAP}_{t+h}/\bar{y}^{GAP}_t$ is the gap ratio. The four channels sum to $\Delta\rho$ exactly.

Each channel answers a distinct economic question. *Within* asks how much aggregate persistence increased because groups' own persistence increased. *Compositional* asks how much aggregate persistence increased because population shares shifted toward groups with higher persistence. *Inequality* asks how much aggregate persistence increased because income dispersion shifted toward higher-persistence groups. *Group gap* asks how much aggregate persistence

increased because between-group wage gaps widened. This within channel isolates changes in the subgroup coefficients. It is not the change in the subgroup component, which also includes the compositional and inequality channels. The identity closes with no residual. Together, the compositional, inequality and group-gap channels are the *between* channels, everything other than a change in the groups' own persistence. Throughout the paper "between" refers to this sum of three channels, and "gap" only to the third.

The identity rests on one choice. When a weight and a rate both move, (8) credits the rate's change at the starting-year weight and the weight's change at the ending-year rate; the mirror choice is equally exact, and the two differ by the cross term $\Delta W \Delta R$. For the three-factor products $n_g \nu_g \rho_g$ the choice is among six orderings of the factors. Rather than pick one, I report the Shapley attribution, which averages each factor's contribution over all orderings [Shorrocks, 2013]. In (9) the weight on $\Delta\rho_g$ becomes $\frac{1}{3}(n_{g,t_0}\nu_{g,t_0} + n_{g,t_1}\nu_{g,t_1}) + \frac{1}{6}(n_{g,t_1}\nu_{g,t_0} + n_{g,t_0}\nu_{g,t_1})$, the weights on Δn_g and on $\Delta\nu_g$ are the same averages of $\nu_g\rho_g$ and of $n_g\rho_g$, and the gap channel becomes $\bar{\omega}^{GAP}\Delta r^{GAP} + \Delta\omega^{GAP}\,\bar{r}^{GAP}$, bars denoting the average of the two dates. The Shapley attribution is symmetric in the two dates, credits nothing to a factor that does not move, and still sums to $\Delta\rho$ exactly. For the gender partition it differs from (9) by a few thousandths; for fine joint partitions, where population shares and variance shares move together, the split of the re-weighting between the compositional and the inequality channel depends on the rule, and their sum is less sensitive to the rule. Appendix F reports the four channels under both pivots, the midpoint and the Shapley rule.

The same construction goes through for the correlation-coefficient measure, starting from its level decomposition in equation (7), with the within-group rates and the weights adjusted accordingly; the details are in Appendix B. The general-K versions of both propositions are derived in Appendix A.

4 Empirics: Application to the Portuguese Data

Population mobility, 1986–2019. Figure 1 plots the population five-year rank-rank slope $\rho_{t,5}$ for base years 1986–2014 (each base year t uses incomes at t and $t+5$, so the last base year draws on 2019 data). Population persistence stands at 0.720 in 1986, dips to its minimum of 0.713 in 1989, rises steadily to a peak of about 0.810 in 2010–2012, and eases to 0.785 by 2014. Over the full window five-year persistence rose by +0.065, a substantial decline in mobility. Relative to its 1986 level, persistence is 12.6% higher at the 2010–2012 peak and remains 9.1% higher in 2014. In cross-country terms this is a marked shift. On the measure used by the Global Repository of Income Dynamics, the expected rank five years ahead of a worker aged 25–55 who starts below the median, Portugal stood in 1987 and 1997 at the level of the most mobile GRID economies, the Nordic countries; by 2007 it sat in the lower third of the comparison, below Germany and Mexico and just above France (Appendix H, Figure 35; Guvenen et al., 2022). Published comparisons on household income and on men's wages likewise place Portugal among the least mobile economies in Europe by the 2000s [Sologon and O'Donoghue, 2009, Van Kerm and Pi Alperin, 2013, Alves and Martins, 2012, OECD, 2018].[14]

[14]The GRID comparison is of time paths rather than levels. GRID ranks workers aged 25–55 on a three-year average of annual earnings, whereas the Portuguese series ranks the same age group on a three-year average of the log hourly wage, so levels are not strictly comparable. Appendix H gives the construction.

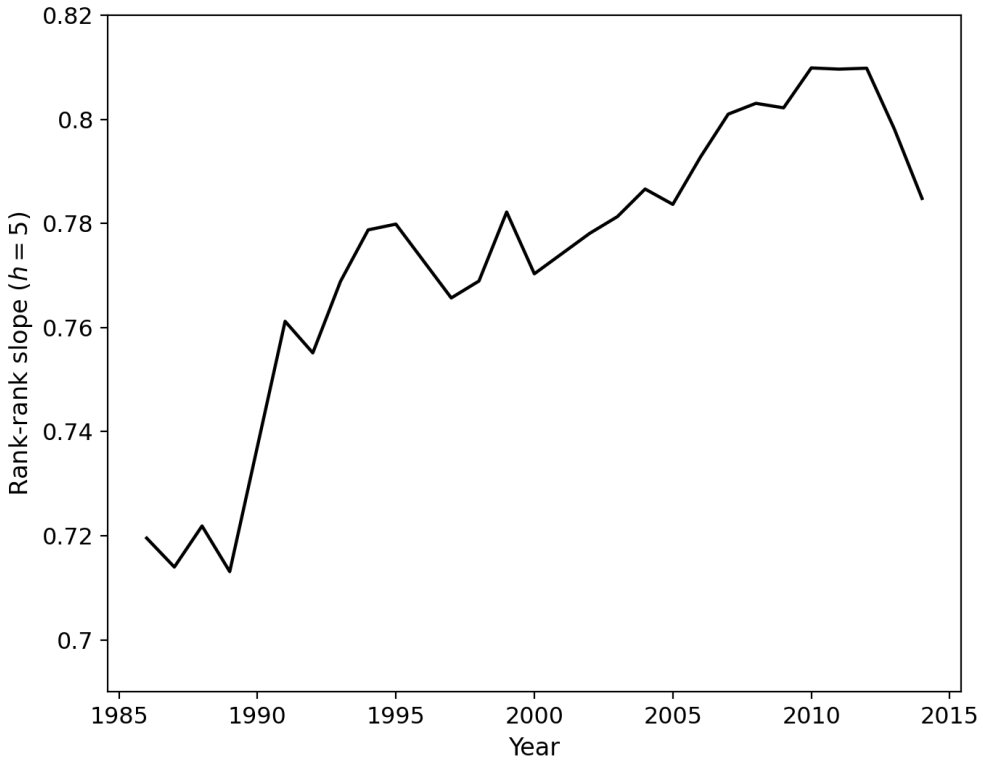


Figure 1: Rank-rank slope (population), $h = 5$, 1986–2014

Notes: Slope of the regression of the within-year percentile rank of the real wage at $t + 5$ on the rank at t, computed on the sample of workers observed at both dates. Base years $t =$ 1986–2014. Higher values indicate more persistence and less mobility. The survey was not conducted in 1990 and 2001, so base years 1990, 1996 and 2001 have no five-year pair; the line is linearly interpolated across them.

The decline is not an artifact of transitory wage fluctuations. Appendix E.3 repeats the exercise for "permanent income" (the centred three-year mean of the log real wage) and finds the hump, if anything, more pronounced (Figure 16).

The same picture emerges from the correlation coefficient. Figure 2 plots the population five-year correlation of log wages, which traces out the same hump; as equation (2) implies, for a given income measure the slope and correlation differ only by the ratio of base- to end-period dispersion, $r = \rho\, \sigma_t / \sigma_{t+h}$, and this ratio moves little over the sample.

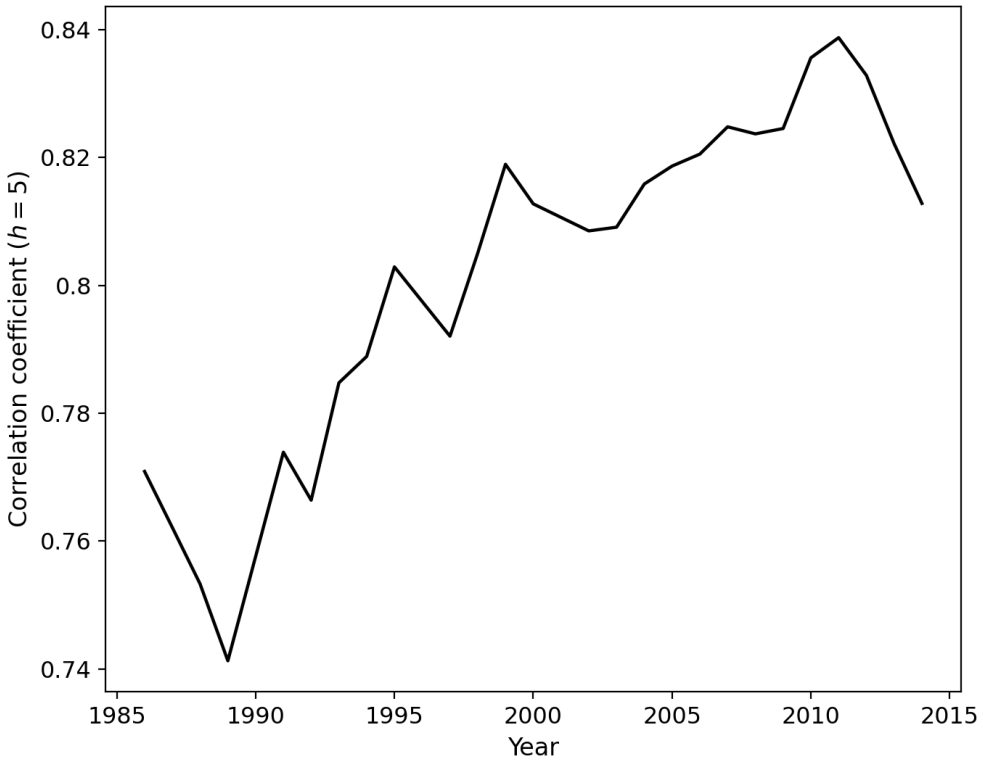


Figure 2: Correlation coefficient (population), $h = 5$

Notes: Pearson correlation between the log real wage at $t + 5$ and at t, computed on the sample of workers observed at both dates; base years 1986–2014. Base years 1990, 1996 and 2001 have no five-year pair (no survey in 1990 and 2001) and are linearly interpolated.

4.1 Applying the level decomposition

Table 1 reports the level decomposition for all five characteristics at once. For each group it lists the within-group slope ρ_g, the population share n_g and the weight ω_g in 1986 and in 2014, together with the population slope and the gap weight ω^{GAP} of each partition. Table 2 and Appendix E.1 report the level decompositions at selected base years; the time paths are shown in the stylized-facts figures of Appendix G.

Table 1: Level decomposition across the five characteristics (rank-rank slope, $h = 5$)

		Persistence ρ_g		Share n_g		Weight ω_g	
Partition	Group	1986	2014	1986	2014	1986	2014
Gender		ρ: 0.720 → 0.785		ω^{GAP}: 0.078 → 0.035			
	Men	0.653	0.736	0.68	0.54	0.610	0.493
	Women	0.762	0.810	0.32	0.46	0.313	0.472
Age		ρ: 0.720 → 0.785		ω^{GAP}: 0.130 → 0.029			
	≤ 31	0.660	0.659	0.41	0.27	0.381	0.217
	32–43	0.787	0.820	0.36	0.42	0.297	0.409
	≥ 44	0.797	0.850	0.23	0.32	0.192	0.346
Education		ρ: 0.718 → 0.785		ω^{GAP}: 0.145 → 0.206			
	Basic	0.688	0.740	0.79	0.26	0.721	0.200
	Secondary	0.594	0.710	0.18	0.53	0.128	0.471
	Higher	0.640	0.703	0.02	0.21	0.005	0.122
Geography		ρ: 0.720 → 0.785		ω^{GAP}: 0.110 → 0.058			
	Lisbon	0.703	0.769	0.34	0.30	0.277	0.292
	North	0.716	0.802	0.40	0.37	0.366	0.350
	Centre–South	0.672	0.738	0.26	0.34	0.246	0.299
Industry		ρ: 0.720 → 0.785		ω^{GAP}: 0.067 → 0.010			
	Goods	0.708	0.801	0.61	0.35	0.583	0.321
	Services	0.714	0.772	0.39	0.65	0.351	0.669

Notes: Decomposition of equation (6) at base years 1986 and 2014. For each partition the header row gives the population slope ρ and the gap weight $\omega^{GAP} = 1 - \sum_g \omega_g$; the group rows give the within-group slope ρ_g, the population share n_g among workers observed at t and $t+5$, and the weight $\omega_g = n_g v_g$. Region: Lisbon, North and a residual Centre–South; industry: goods (sections A–F) and services (G–U). The education partition drops workers with unrecorded schooling.

Gender. Women's wages are more persistent than men's in every year of the sample (Figure 21, bottom-left panel), 0.810 against 0.736 in 2014. The gender wage gap, though shrinking, still separates the two groups. The gap term accounts for 0.083 of the population slope in 1986 and 0.040 in 2014. Women's share of the workforce rose from 0.31 to 0.46, and their weight from 0.31 to 0.47.

Age. Older workers' wages are far more persistent than younger workers', 0.850 for the oldest tercile against 0.659 for the youngest in 2014, and the ordering holds in every year. Only the middle and oldest terciles became stickier over time; the youngest ended essentially where it started (0.660 in 1986). As the workforce aged, the oldest tercile's share rose from 0.23 to 0.32 and its weight from 0.19 to 0.35.

Education. The basic-educated are the least mobile group at the start and at the end of the sample, 0.740 against 0.710 for secondary and 0.703 for higher education in 2014, but the ordering is not stable over time. In 2007–2011 the higher-educated were the most persistent group (Figure 27). What holds in every year is that all three groups sit below the population slope, the configuration of Section 3.1.1. The education wage gap keeps workers from crossing between education groups. The basic-educated share collapsed from 0.79 to 0.26 as the workforce upskilled, and its weight from 0.72 to 0.20.

Region. Workers in the Centre–South are the most mobile in every year (0.738 in 2014, against 0.769 in Lisbon and 0.802 in the North); Lisbon and the North are the two persistent regions and trade the top spot over time. Regional wage gaps are modest, so the gap term is small and shrinks, from 0.097 in 1986 to 0.058 in 2014. Employment moved away from Lisbon (0.34 to 0.30) and the North (0.40 to 0.37) toward the Centre–South (0.26 to 0.34).

Industry. Goods and services differ little in persistence (0.801 against 0.772 in 2014), and which sector is the more persistent changes from year to year. The gap term is small throughout and falls from 0.057 to 0.012 as the sectoral wage gap narrows. Employment moved into services, whose share rose from 0.39 to 0.65 and whose weight from 0.35 to 0.67.

Two features of Table 1 matter for what follows. First, the gap term is large only where between-group wage gaps are large, that is, for education and age; it is small for region and industry, whose groups overlap in the wage distribution. Second, the weights move with the share of the wage variance a group owns and not only with its headcount. The oldest tercile's

weight rose by almost twice as much as its share, because the dispersion of wages among older workers rose relative to the population's.

Great Gatsby curves. The cross-country Great Gatsby curve links higher inequality to lower income mobility [Corak, 2013, Guvenen et al., 2022]. Within Portugal the relationship can be drawn in two ways (Appendix H). Across subgroups it holds. Pooling every group of every partition in every base year, more unequal subgroups are less mobile, with a correlation of 0.51 between a subgroup's dispersion of log wages and its own rank-rank slope, and of 0.62 once subgroup-years are weighted by population. The gradient is steep for age and present for region and industry; it is flat for education on the rank-rank slope, though positive on the log-wage elasticity; and gender runs the other way, more dispersed men being if anything slightly more mobile than less dispersed women. Across periods, by contrast, there is no Great Gatsby relationship. Between base years, the dispersion of log wages and five-year persistence are essentially uncorrelated (0.23 unweighted, 0.06 weighted by sample size). Appendix H.1 shows how the two can come apart. Inequality and immobility move together only when the additional dispersion is permanent.

4.2 Applying the changes decomposition

The main finding is that income mobility within finely defined groups barely changed. No single grouping accounts for the decline. Cut by gender, age, education, region or sector alone, the within channel remains large, never below two fifths of the change. Only when workers are grouped jointly by age and education does the within channel collapse, to +0.004 of a +0.067 change by 2014, and the decline is accounted for by the between channels, the re-weighting of the labor force across age–education cells and the persistence of the wage gaps between them. Throughout, "between" denotes the sum of the three channels other than within (compositional, inequality and group-gap), and every attribution uses the Shapley rule of Section 3.2.

Gender. Table 2 reports the level decomposition (6) at selected base years. Table 3 reports the four-channel decomposition (9) of $\Delta\rho$ from the 1986 base year to selected end years, and Figure 3 plots the four channels year by year. Under this partition the within channel accounts for the whole of the decline in mobility. From 1986 to 2014 the population slope rose +0.065; the within channel alone contributes +0.065. The compositional channel, feminization, the shift of population weight toward the more persistent group, contributes +0.024, and the inequality channel +0.020. The group-gap channel is mobility-*enhancing* at −0.044. The compression of the gender wage gap dismantled part of the barrier between the two groups, and almost exactly offsets the two re-weighting channels, so that the between channels net to approximately zero. At the 2010–2012 peak the same pattern holds with larger magnitudes ($\Delta\rho \approx +0.090$, within $\approx +0.091$).

Table 2: Level decomposition by gender (rank-rank slope, $h = 5$)

Year	ρ	Subgroup	ρ_M	ρ_F	ω_M	ω_F	ω^{GAP}	r^{GAP}
1986	0.720	0.636	0.653	0.762	0.610	0.313	0.078	1.072
1994	0.779	0.694	0.724	0.792	0.550	0.374	0.076	1.109
2004	0.787	0.736	0.733	0.824	0.527	0.425	0.048	1.049
2011	0.810	0.771	0.766	0.834	0.492	0.473	0.036	1.081
2014	0.785	0.745	0.736	0.810	0.493	0.472	0.035	1.120

Notes: Decomposition of equation (6). "Subgroup" is $\omega_M\rho_M + \omega_W\rho_W$; the gap term has weight $\omega^{GAP} = 1 - \omega_M - \omega_W$ and gap ratio $r^{GAP} = \bar{y}^{GAP}_{t+h} / \bar{y}^{GAP}_{t} = (\bar{y}_{M,t+h} - \bar{y}_{W,t+h})/(\bar{y}_{M,t} - \bar{y}_{W,t})$, so that $\rho = \text{Subgroup} + \omega^{GAP} r^{GAP}$ closes exactly. Ranks are within-year population percentile ranks of the real wage.

Table 3: Four-channel change decomposition (gender partition)

Year t	$\Delta\rho$	Within	Compositional	Inequality	Group-gap
1991	+0.0416	+0.0382	+0.0062	-0.0137	+0.0109
1994	+0.0592	+0.0515	+0.0102	-0.0037	+0.0012
2004	+0.0671	+0.0681	+0.0180	+0.0139	-0.0329
2011	+0.0901	+0.0908	+0.0241	+0.0200	-0.0448
2014	+0.0652	+0.0649	+0.0239	+0.0201	-0.0437

Notes: Decomposition of equation (9) from base year $t_0 = 1986$ to the indicated end year t_1 (each at horizon $h = 5$), Shapley attribution. The four channels sum to $\Delta\rho$ exactly, up to rounding.

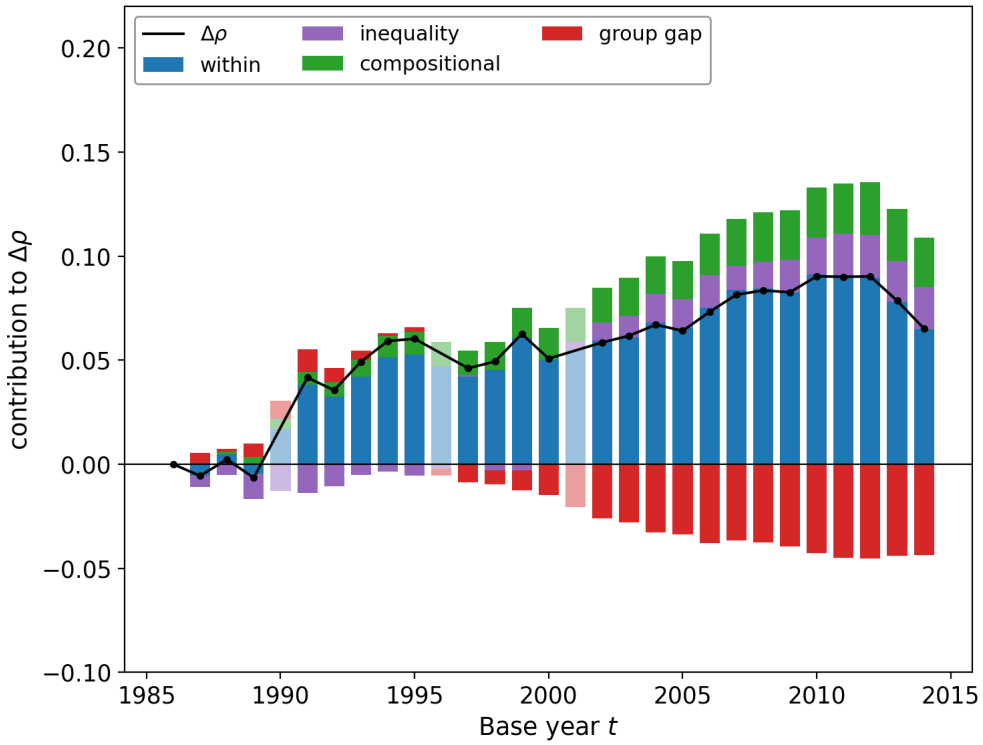


Figure 3: Change decomposition relative to 1986 (gender)

Notes: Cumulative contribution of each channel of equation (9) to $\Delta\rho$ relative to the 1986 base year, Shapley attribution. The line is the population change; the bars are the four channels. Base years 1990, 1996 and 2001 have no five-year pair (no survey in 1990 and 2001); their bars are linear interpolations, drawn lighter.

The negative red bars from around 1997 onward indicate that the group-gap channel contributes negatively to the change in persistence, reflecting that the gap between men and women has been eroding as the gender pay gap compresses, visible as the narrowing distance between the male and female lines in the mean-wage (μ_t) panel of Figure 21. Strictly, this channel tracks the *persistence* of the gap. The weight $\omega^{GAP} \propto (\bar{y}_t^{GAP})^2$ falls as the gap narrows, so a shrinking gender gap contributes to the negative bar. Most of the increase in population persistence since 1990, however, is driven by rising persistence *within* each subgroup, shown by the dominant blue "within" bars. Both men and women became individually more persistent (ρ_W rose from 0.76 to a peak of 0.83 and ρ_M from 0.65 to 0.77), as both lines drift upward in the persistence (ρ_t) panel of Figure 21. The green "compositional" bars contribute positively throughout, because population weight shifted toward the higher-persistence group.

Women's share of the workforce grew from 32% to 46%, traced by the rising female line in the composition (n_t) panel of Figure 21, and women are the stickier group. Finally, the purple "inequality" bars start negative and turn positive around 2001–2002. This channel measures how each group's variance share, the variance-share (v_t) panel of Figure 21, moves *relative to its own 1986 level*, weighted by population share and persistence. Because women (the higher-persistence group) saw their variance share fall *below* its 1986 value through the 1990s and then rise *above* it from about 2002, the re-weighting first pulled the aggregate toward the lower-persistence group (negative contribution) and later toward the higher-persistence group (positive contribution). This is a within-group change over time, not a comparison between men and women. Women's variance share is in fact *higher* than men's in every year of the sample.

The other four one-dimensional partitions are reported in Appendix E.1 (Figures 6–9). Each still carries a sizable within channel. Region (+0.068) and industry (+0.071) behave like gender and attribute the whole change to within-group persistence. Education attributes nine tenths of it to within (+0.060 of +0.067), the two re-weighting channels and the group-gap channel netting out. Age is the only one-dimensional partition in which the between channels do substantial work. Its within channel is +0.025, two fifths of the change, because the aging of the workforce moved weight toward its older and more persistent members. A single characteristic never removes the within channel; a pair does.

The age×education partition. Table 4 and Figure 4 report the decomposition for the nine cells formed by crossing the three age terciles with the three education groups. The within channel is +0.004 by 2014, six percent of the +0.067 change, and +0.024 at the 2011 peak, a quarter of it. Once workers are compared with others of the same age and education, churning inside cells did not slow over three decades. What changed is the structure the churning happens in. The two re-weighting channels sum to +0.050 (compositional −0.044, inequality +0.094) and the group-gap channel adds +0.013. Together these between channels account

for the rise in persistence.

Table 4: Four-channel change decomposition (age×education partition, $K = 9$)

Year t	$\Delta\rho$	Within	Compositional	Inequality	Group-gap
1991	+0.0434	+0.0138	-0.0167	-0.0147	+0.0610
1994	+0.0529	+0.0214	-0.0102	+0.0188	+0.0229
2004	+0.0694	+0.0223	-0.0331	+0.0558	+0.0245
2011	+0.0917	+0.0244	-0.0424	+0.0768	+0.0329
2014	+0.0667	+0.0039	-0.0437	+0.0936	+0.0129

Notes: Decomposition of equation (9) for the joint age×education partition (three age terciles crossed with basic, secondary and higher education), from base year $t_0 = 1986$ to the indicated end year, horizon $h = 5$, Shapley attribution. Ranks are population percentile ranks over all workers; a worker enters a base year when base-year education is recorded, so the $\Delta\rho$ column coincides with the education partition's. Education is unrecorded in 1995, 1997 and 1998, so those base years drop out.

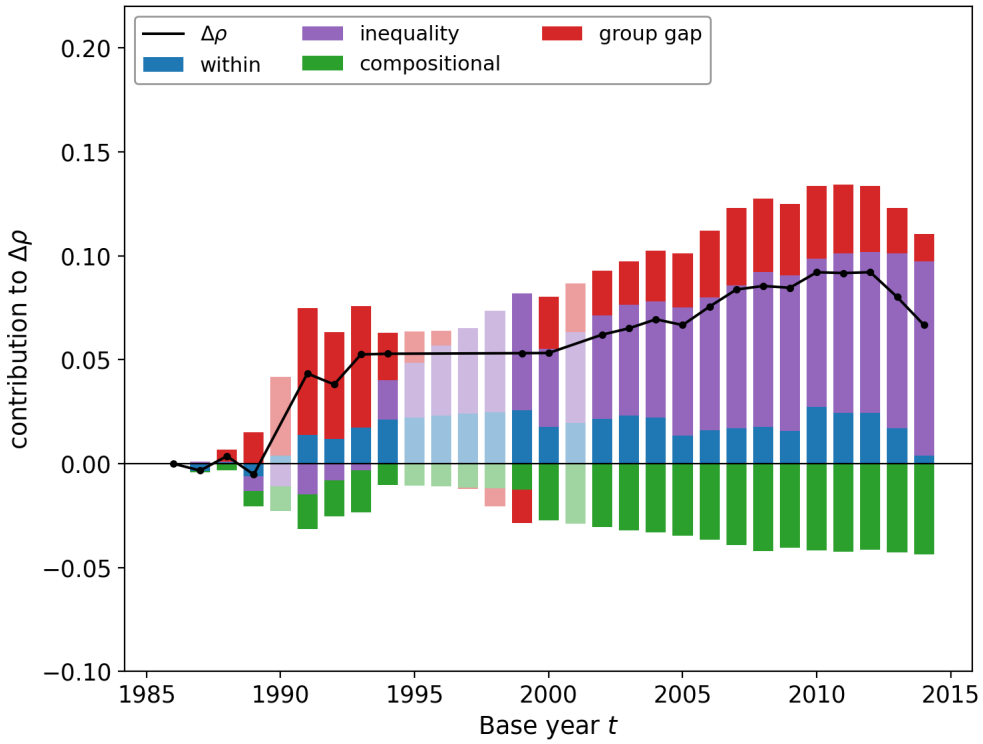


Figure 4: Change decomposition relative to 1986 (age×education)

Notes: Cumulative contribution of each channel of equation (9) to $\Delta\rho$ relative to the 1986 base year, age×education partition ($K = 9$), Shapley attribution. The line is the population change; the bars are the four channels. Base years without a five-year pair or without recorded education (1990, 1995–1998, 2001) are linearly interpolated and drawn lighter.

Age×education is special. Table 15 in Appendix E.2 reports the decomposition for all ten pairs of characteristics. No other pair removes the within channel. The smallest within con-

tributions among the remaining pairs are +0.028 for gender×age, +0.030 for age×region and +0.031 for age×industry, each above two fifths of the change, and the pairs that do not involve age leave between +0.046 and +0.069 to within. Adding gender as a third dimension (eighteen cells) confirms the result, with a within channel of −0.011. Every pair that involves age cuts the within channel to about half, because the age partition alone already attributes three fifths of the change to between-group forces; education removes what remains.

Which dimension does the work? Because the between channels of the age×education partition absorb what the age-only and education-only partitions attribute to their within channels, the natural accounting is a Shapley decomposition over the two dimensions, on the common education-complete sample. Each dimension is credited with its own between contribution plus half of what the two add jointly (Appendix E.2, Table 16). Of the +0.063 between-cell change by 2014, +0.048 is attributable to age and +0.015 to education. The two dimensions work through different channels and partly against each other. Aging re-weights the workforce toward the older, more persistent cells (a re-weighting contribution of +0.11 attributable to age), while the gaps between age groups compressed as each cohort of young workers caught up (gap contribution −0.06). Upskilling re-weights away from the persistent basic-education cells (−0.06), while the education wage gaps themselves became more important (+0.08). Splitting the gap channel exactly by pair type, pairs of cells that differ in education contribute +0.063 and pairs that differ only in age −0.050. The rise in persistence is therefore neither the skill premium alone nor aging alone. It is the combination of an aging workforce, which puts ever more weight on the most persistent cells, and of persistent education wage gaps, net of the mobility-enhancing effects of upskilling and of narrowing age gaps.

Two cautions apply to the between channels. First, within the re-weighting pair, the split between the compositional and the inequality channel is not robust to the attribution rule when population shares and variance shares move together, as they do for the fine cells. The pair's sum is less sensitive to the four rules of Appendix F than its components. The text therefore reasons with the sum. Second, the within channel of the age×education partition

ranges from −0.017 (base pivot) to +0.028 (inverted pivot) across attribution rules, with the symmetric rules at +0.004 and +0.006; under every rule it is below the within channel of any one-dimensional partition, so the conclusion does not depend on the rule.

The correlation of log wages. Repeating the exercise for the correlation of log wages, the natural measure when income enters in logs (Section 2.1), delivers the same verdict (Appendix E.3, Table 17). The within channel carries the whole change under gender, region and industry, about half of it under age, and vanishes under age×education (−0.001 of +0.044). The two measures differ in how they attribute the rest, in part because the correlation weights groups by their share of log-wage dispersion, which moved across groups far more than rank variance did; under education the correlation's within channel is +0.024 of +0.044 where the slope's is +0.060 of +0.067. Appendix E.3 discusses these differences.

5 Discussion: the level of churning and the content of the decomposition

Taken as a whole, the decompositions tell a consistent story about the Portuguese mobility decline, and a cautionary one about how to read it. Across every one-dimensional partition except age, the rise in population persistence is largely a *within*-group phenomenon. Men, women, all three education groups, and the two older age terciles each became individually less mobile, and under gender, region and industry the within channel accounts for the whole change. The changes in the workforce's composition that dominate public discussion of Portuguese labor-market change, feminization, upskilling and the move into services, contribute little on their own; aging is the exception, and it works through the re-weighting of the workforce toward its most persistent members rather than through headcounts alone. The compression of the gender wage gap worked to *enhance* mobility; the persistence of the education wage gaps worked against it. Read one characteristic at a time, then, composition is not the story.

The joint age×education decomposition overturns this reading. Once workers are grouped by both age and education, the within channel vanishes, and the between channels, the re-weighting of the workforce across age–education cells and the gaps between them, carry the rise in persistence. This is not a contradiction. It is the same decomposition computed on smaller groups, and smaller groups leave less room for churning within them, so more of the same aggregate change gets attributed to differences *between* cells rather than to movement inside them. But putting the coarse and fine cuts side by side raises a question the one-dimensional results leave implicit. At what level does churning actually happen? There are good reasons to prefer the finer cell, people with similar observable characteristics compete for and circulate through similar jobs, so a young graduate realistically churns against other young graduates, not against a sixty-year-old with basic schooling. Yet the logic has no natural stopping point. One can keep adding dimensions, occupation, region, firm, tenure, until almost nothing is left inside the cells to move the aggregate at all. The choice of partition is, in that sense, somewhat arbitrary, and the "within versus between" split should be read as conditional on it.

This points to what the framework is, and what it is not. The decompositions are exact algebraic identities, disciplined ways of accounting for where a change in a population statistic comes from, given a partition, and they are genuinely useful as description. They tell us that Portuguese income mobility fell because the workforce was re-weighted toward its more persistent age–education cells and because the wage gaps between those cells persisted, while churning within cells did not slow. But an accounting identity carries no inherent economic or causal content. It does not explain why the education premium widened, why the workforce aged, or whether any of these forces would survive a change in policy or institutions; nor does it privilege one partition as the "true" level of churning. Answering those questions requires a structural model of how workers sort across jobs and how wages are set, which is beyond the scope of this paper.

6 Conclusion

Population income mobility hides substantial heterogeneity across population subgroups, and this paper provides the accounting that connects the two. The level of any slope-based mobility measure, regression slope or correlation coefficient, decomposes exactly into within-group mobilities, weighted by composition *times* inequality, plus the persistence of the wage gaps between groups; population mobility is not a weighted average of subgroup mobilities, and can be lower than every one of them. Changes in population mobility decompose exactly into four channels, within, compositional, inequality, and group gap, in a single closed-form identity that requires no counterfactual simulations. The framework applies to intragenerational and intergenerational measures alike.

Applied to the universe of Portuguese private-sector workers over 1986–2019, the decomposition shows that no single characteristic accounts for the decline in income mobility. Grouped by gender, region, industry or education alone, workers appear to have become individually less mobile, and only age, through the aging of the workforce, attributes a substantial share of the change to forces between groups. Grouped jointly by age and education, the picture changes. Churning within age–education cells did not slow at all, and the rise in persistence is accounted for by the re-weighting of the workforce toward its older and more persistent cells and by the persistence of education wage gaps, net of the mobility-enhancing effects of upskilling and of narrowing age and gender gaps. The explanation of Portugal's mobility decline is therefore to be sought in the structure of its labor force, the age composition of the workforce and the education wage gaps it sustains, rather than in a general stickiness of individual wages.

References

N. Alves and C. Martins. Mobility and income inequality in the European Union and in Portugal. *Banco de Portugal Economic Bulletin*, Summer:57–70, 2012.

P. J. Bickel, E. A. Hammel, and J. W. O'Connell. Sex bias in graduate admissions: Data from Berkeley. *Science*, 187(4175):398–404, 1975.

M. Buchinsky, G. Fields, D. Fougère, and F. Kramarz. Francs or ranks? Earnings mobility in France, 1967–1999. Cepr discussion paper, CEPR, 2003. URL https://papers.ssrn.com/sol3/papers.cfm?abstract_id=432721.

A. R. Cardoso. Wage mobility: Do institutions make a difference? *Labour Economics*, 13(3): 387–404, 2006.

A. Carneiro, P. Portugal, P. Raposo, and P. M. M. Rodrigues. The persistence of wages. *Journal of Econometrics*, 233(2):596–611, 2023.

R. Chetty, N. Hendren, P. Kline, and E. Saez. Where is the land of opportunity? the geography of intergenerational mobility in the united states. *Quarterly Journal of Economics*, 129(4): 1553–1623, 2014.

R. Chetty, N. Hendren, M. R. Jones, and S. R. Porter. Race and economic opportunity in the United States: An intergenerational perspective. *Quarterly Journal of Economics*, 135(2): 711–783, 2020. doi: 10.1093/qje/qjz042.

M. Corak. Income inequality, equality of opportunity, and intergenerational mobility. *Journal of Economic Perspectives*, 27(3):79–102, 2013.

F. A. Cowell and E. Flachaire. Measuring mobility. *Quantitative Economics*, 9(2):865–901, 2018. doi: 10.3982/QE512.

F. A. Cowell and E. Flachaire. Mobility and mobility measures. Working Paper 2502.19608, arXiv, 2025. URL https://arxiv.org/abs/2502.19608.

F. Guvenen, L. Pistaferri, and G. L. Violante. Global trends in income inequality and income dynamics: New insights from grid. *Quantitative Economics*, 13(4):1321–1360, 2022.

T. Hertz. A group-specific measure of intergenerational persistence. *Economics Letters*, 100(3): 415–417, 2008. doi: 10.1016/j.econlet.2008.03.008.

E. Jácome, I. Kuziemko, and S. Naidu. Mobility for all: Representative intergenerational mobility estimates over the twentieth century. *Journal of Political Economy*, 133(1), 2025.

M. Jäntti and S. P. Jenkins. Income mobility. In A. B. Atkinson and F. Bourguignon, editors, *Handbook of Income Distribution*, volume 2, chapter 10, pages 807–935. Elsevier, 2015. doi: 10.1016/B978-0-444-59428-0.00011-4.

K. B. Karlson and J. In. Decomposing sibling correlations: a new measure of group-specific intergenerational persistence in socioeconomic outcomes. *European Sociological Review*, 41 (5):822–830, 2025. doi: 10.1093/esr/jcae035.

C.-I. Lee and G. Solon. Trends in intergenerational income mobility. NBER Working Paper 12007, National Bureau of Economic Research, 2006.

R. McGee. Sufficient statistics for economic mobility: When do measures agree? IFS Working Paper 26/38, Institute for Fiscal Studies, 2026. URL https://doi.org/10.1920/wp.ifs.2026.3826.

OECD. *A Broken Social Elevator? How to Promote Social Mobility*. OECD Publishing, Paris, 2018. doi: 10.1787/9789264301085-en.

C. Oliveira. Income and wage inequality in democratic Portugal, 1974–2020. *Fiscal Studies*, 45(3):393–414, 2024. doi: 10.1111/1475-5890.12391.

A. F. Shorrocks. The class of additively decomposable inequality measures. *Econometrica*, 48 (3):613–625, 1980. doi: 10.2307/1913126.

A. F. Shorrocks. Decomposition procedures for distributional analysis: a unified framework based on the Shapley value. *Journal of Economic Inequality*, 11(1):99–126, 2013.

D. M. Sologon and C. O'Donoghue. Increased opportunity to move up the economic ladder? Earnings mobility in EU: 1994–2001. IZA Discussion Paper 4311, IZA, Bonn, 2009.

D. M. Sologon and C. O'Donoghue. Earnings mobility, earnings inequality, and labor market institutions in Europe. In *Inequality, Mobility and Segregation: Essays in Honor of Jacques Silber*, volume 20 of *Research on Economic Inequality*, pages 237–283. Emerald, Bingley, 2012.

F. Torche and A. Corvalán. Estimating intergenerational mobility with grouped data: A critique of clark's the son also rises. *Sociological Methods & Research*, 47(4):787–811, 2018.

P. Van Kerm and M. N. Pi Alperin. Inequality, growth and mobility: The intertemporal distribution of income in European countries 2003–2007. *Economic Modelling*, 35:931–939, 2013.

A The general K-group decompositions

This appendix derives the level and changes decompositions of the slope for an arbitrary number of groups K; the correlation versions are in Appendix B. Partition the population into K mutually exclusive groups $g = 1, \ldots, K$, with $N = \sum_g N_g$ and shares $n_g = N_g/N$. For a pair (g, k) define the between-group mean gap $\bar{y}_s^{GAP_{gk}} = \bar{y}_{g,s} - \bar{y}_{k,s}$, $s \in \{t, t+h\}$.

A.1 Level

Partitioning the numerator of (3) over the K groups and adding and subtracting group means exactly as in Section 3.1 gives, for each group,

$$\sum_{i \in g} (y_{i,t+h} - \bar{y}_{t+h})(y_{i,t} - \bar{y}_t) = N_g \, \mathrm{Cov}_g + N_g (\bar{y}_{g,t+h} - \bar{y}_{t+h})(\bar{y}_{g,t} - \bar{y}_t),$$

so that summing over g,

$$\sum_i (y_{i,t+h} - \bar{y}_{t+h})(y_{i,t} - \bar{y}_t) = \sum_{g=1}^{K} N_g \, \mathrm{Cov}_g + \underbrace{\sum_{g=1}^{K} N_g (\bar{y}_{g,t+h} - \bar{y}_{t+h})(\bar{y}_{g,t} - \bar{y}_t)}_{\text{between term}}. \tag{10}$$

The between term converts into a sum over group pairs by the weighted-covariance identity: for any group-level values a_g, b_g with weighted means $\bar{a} = \frac{1}{N}\sum_g N_g a_g$, $\bar{b} = \frac{1}{N}\sum_g N_g b_g$,

$$\sum_{g=1}^{K} N_g (a_g - \bar{a})(b_g - \bar{b}) = \frac{1}{N} \sum_{g=1}^{K-1} \sum_{k=g+1}^{K} N_g N_k (a_g - a_k)(b_g - b_k). \tag{11}$$

Proof of (11). Both sides equal $\sum_g N_g a_g b_g - N \bar{a} \bar{b}$. For the left side this is immediate by expanding and using $\sum_g N_g a_g = N \bar{a}$. For the right side, write the pair sum as $\frac{1}{2}\sum_g \sum_k$; expanding $(a_g - a_k)(b_g - b_k)$ and summing gives $2N \sum_g N_g a_g b_g - 2N^2 \bar{a} \bar{b}$; multiplying by $\frac{1}{2N}$ yields the claim. □

Applying (11) with $a_g = \bar{y}_{g,t+h}$, $b_g = \bar{y}_{g,t}$, the numerator becomes

$$\sum_{g=1}^{K} N_g \,\mathrm{Cov}_g + \sum_{g<k} \frac{N_g N_k}{N}\, \bar{y}_{t+h}^{\,GAP_{gk}}\, \bar{y}_{t}^{\,GAP_{gk}}, \tag{12}$$

and the denominator (the same object with $y_{i,t+h} \to y_{i,t}$) becomes

$$N\ \mathrm{Var}(y_{i,t}) = \sum_{g=1}^{K} N_g \sigma_{g,t}^2 + \sum_{g<k} \frac{N_g N_k}{N} \big(\bar{y}_{t}^{\,GAP_{gk}}\big)^2. \tag{13}$$

Dividing term by term as in Section 3.1:

$$\boxed{\rho_{t,h} = \sum_{g=1}^{K} \omega_g\, \rho_g + \sum_{g=1}^{K-1} \sum_{k=g+1}^{K} \omega^{GAP_{gk}}\, \frac{\bar{y}_{t+h}^{\,GAP_{gk}}}{\bar{y}_{t}^{\,GAP_{gk}}}} \tag{14}$$

with

$$\omega_g = n_g\, \frac{\sigma_{g,t}^2}{\sigma_t^2}, \qquad \omega^{GAP_{gk}} = n_g n_k\, \frac{(\bar{y}_t^{\,GAP_{gk}})^2}{\sigma_t^2}, \qquad \sum_{g} \omega_g + \sum_{g<k} \omega^{GAP_{gk}} = 1.$$

Setting $K = 2$ recovers Proposition 1. The pairwise form of the gap term is essential for interpretation, it identifies *which* group gaps carry it, and for the changes decomposition below.

A.2 Changes

From (14), population mobility at each date is $\rho = \sum_a W_a R_a$, a sum over K weighted subgroup contributions and $\binom{K}{2}$ gap pairs. Differencing between base years t_0 and t_1 and applying the product–difference identity (8) to every term, then splitting each subgroup weight change via

$\Delta\omega_g = \Delta(n_g v_g) = \Delta n_g\, v_{g,t_1} + \Delta v_g\, n_{g,t_0}$, gives

$$\Delta\rho = \underbrace{\sum_{g=1}^{K} \omega_{g,t_0}\,\Delta\rho_g}_{\text{within}} + \underbrace{\sum_{g=1}^{K} \Delta n_g\, v_{g,t_1}\,\rho_{g,t_1}}_{\text{compositional}} + \underbrace{\sum_{g=1}^{K} \Delta v_g\, n_{g,t_0}\,\rho_{g,t_1}}_{\text{inequality}} + \underbrace{\sum_{g=1}^{K-1}\sum_{k=g+1}^{K} \left[\omega_{t_0}^{GAP_{gk}}\,\Delta r^{GAP_{gk}} + \Delta\omega^{GAP_{gk}}\, r_{t_1}^{GAP_{gk}}\right]}_{\text{group gap}}, \tag{15}$$

where $r^{GAP_{gk}} = \bar{y}_{t+h}^{GAP_{gk}} / \bar{y}_t^{GAP_{gk}}$. The channels sum to $\Delta\rho$ exactly, since the pieces telescope back to $\rho_{t_1} - \rho_{t_0}$. Setting $K = 2$ recovers Proposition 2. The Shapley attribution of Section 3.2 replaces, group by group and pair by pair, the base-year weights and end-year rates of (15) by their averages over the orderings of the factors; it applies to any K.

B The correlation-coefficient decompositions

The correlation replaces the slope's denominator $\sigma_t\sigma_t$ with the cross-period product $\sigma_t\sigma_{t+h}$:

$$\rho_{t,h} = \frac{\mathrm{Cov}(y_{i,t+h}, y_{i,t})}{\sigma_t\,\sigma_t} \qquad \text{vs.} \qquad r_{t,h} = \frac{\mathrm{Cov}(y_{i,t+h}, y_{i,t})}{\sigma_t\,\sigma_{t+h}}.$$

The numerator development of Appendix A is unchanged, equation (12) still gives the exact split of the covariance into K within pieces and $\binom{K}{2}$ gap pieces. What changes is that the cross-period product $\sigma_t\sigma_{t+h}$, unlike the single variance σ_t^2, has no law-of-total-variance expansion into group pieces. There is no analogue of (13). The three (or $K + \binom{K}{2}$) numerator pieces are therefore divided by $N\sigma_t\sigma_{t+h}$ directly, and each is turned into a weight-times-rate product by

inserting the group's own dispersion:

$$\frac{N_g \operatorname{Cov}_g}{N\,\sigma_t\sigma_{t+h}} = \underbrace{n_g \frac{\sigma_{g,t}\,\sigma_{g,t+h}}{\sigma_t\,\sigma_{t+h}}}_{\omega_g^c} \underbrace{\frac{\operatorname{Cov}_g}{\sigma_{g,t}\,\sigma_{g,t+h}}}_{r_g},$$

$$\frac{(N_g N_k/N)\,\bar{y}_{t+h}^{GAP_{gk}}\,\bar{y}_t^{GAP_{gk}}}{N\,\sigma_t\sigma_{t+h}} = \underbrace{n_g n_k \frac{(\bar{y}_t^{GAP_{gk}})^2}{\sigma_t\,\sigma_{t+h}}}_{\omega^{c,GAP_{gk}}} \underbrace{\frac{\bar{y}_{t+h}^{GAP_{gk}}}{\bar{y}_t^{GAP_{gk}}}}_{\text{gap ratio}}.$$

Summing over groups and pairs gives the level decomposition of the correlation:

$$\boxed{r_{t,h} = \sum_{g=1}^{K} \omega_g^c\, r_g + \sum_{g=1}^{K-1}\sum_{k=g+1}^{K} \omega^{c,GAP_{gk}} \frac{\bar{y}_{t+h}^{GAP_{gk}}}{\bar{y}_t^{GAP_{gk}}}} \tag{16}$$

with

$$\omega_g^c = n_g \frac{\sigma_{g,t}\,\sigma_{g,t+h}}{\sigma_t\,\sigma_{t+h}}, \qquad \omega^{c,GAP_{gk}} = n_g n_k \frac{(\bar{y}_t^{GAP_{gk}})^2}{\sigma_t\,\sigma_{t+h}}, \qquad r_g = \frac{\operatorname{Cov}_g}{\sigma_{g,t}\,\sigma_{g,t+h}}.$$

Setting $K = 2$ gives equation (7) of the main text.

Two remarks. (i) Because the denominator is the cross-period product rather than a single variance, the correlation weights do *not* sum to one, the exact-adding-up property is specific to the slope. (ii) As a consistency check, dividing (14) by σ_{t+h}/σ_t term by term reproduces (16). The whole expression equals the slope times the ratio of base- to end-period standard deviations, $r_{t,h} = \rho_{t,h}\,\sigma_t/\sigma_{t+h}$.

Changes. The changes decomposition is constructed exactly as in Appendix A. Difference $r = \sum_a W_a R_a$ term by term, apply the product–difference identity (8), and split the subgroup weight $\omega_g^c = n_g d_g$, where $d_g = \sigma_{g,t}\sigma_{g,t+h}/(\sigma_t\sigma_{t+h})$ is the *dispersion share*, via $\Delta\omega_g^c = \Delta n_g\, d_{g,t_1} +$

$\Delta d_g\, n_{g,t_0}$:

$$\Delta r = \underbrace{\sum_{g=1}^{K} \omega^{c}_{g,t_0}\, \Delta r_g}_{\text{within}} + \underbrace{\sum_{g=1}^{K} \Delta n_g\, d_{g,t_1}\, r_{g,t_1}}_{\text{compositional}} + \underbrace{\sum_{g=1}^{K} \Delta d_g\, n_{g,t_0}\, r_{g,t_1}}_{\text{inequality}} + \underbrace{\sum_{g=1}^{K-1} \sum_{k=g+1}^{K} \left[\omega^{c,GAP_{gk}}_{t_0}\, \Delta r^{GAP_{gk}} + \Delta \omega^{c,GAP_{gk}}\, r^{GAP_{gk}}_{t_1} \right]}_{\text{group gap}}. \tag{17}$$

The four-channel structure is identical to the slope's; the only change is that the variance share v_g becomes the dispersion share d_g, the correlation's analogue of the variance share, and within slopes become within correlations.

C Churning with changing dispersion

Consider an AR(1) process

$$y_{i,t+h} = \phi\, y_{i,t} + \varepsilon_{i,t+h}, \tag{18}$$

with autoregressive coefficient ϕ, where $\varepsilon_{i,t+h}$ is independent of $y_{i,t}$. The slope of the regression of $y_{i,t+h}$ on $y_{i,t}$ is

$$\rho_{t,h} = \frac{\mathrm{Cov}(y_{i,t+h}, y_{i,t})}{\sigma_t^2} = \phi, \tag{19}$$

while the correlation coefficient is

$$r_{t,h} = \frac{\mathrm{Cov}(y_{i,t+h}, y_{i,t})}{\sigma_t\, \sigma_{t+h}} = \phi\, \frac{\sigma_t}{\sigma_{t+h}}. \tag{20}$$

Now let the variance of the innovation $\varepsilon_{i,t+h}$ rise suddenly. This produces a large increase in churning. Individuals change their positions in the income distribution. Yet the slope is unchanged at ϕ, because the additional reshuffling loads entirely on σ_{t+h}, which the slope's denominator σ_t^2 never sees. The correlation, by contrast, falls, since σ_{t+h} rises while the covari-

ance does not, correctly registering the increase in churning. This is why, for log income, the correlation rather than the slope is the measure that stays at one under unchanged positions and moves with genuine reshuffling.

D Simpson's paradox and the level decomposition

D.1 The classic case

Population mobility can be *lower than the mobility of every subgroup*. Subgroups may exhibit substantial within-group mobility, and yet overall mobility can be small, because the within-group slopes enter with weights that sum to less than one while the residual weight loads onto the persistence of group gaps. This is an application of Simpson's paradox. The slope in each group can be low while the pooled slope is high, because pooling adds the between-group relationship to the within-group ones. In the classic versions of the paradox the reversal is starker, with the pooled relationship taking the *opposite sign* of every within-group relationship, as in the Berkeley admissions data of Bickel et al. [1975], where individual departments favored female applicants while the population appeared to favor men, the population loading on the fact that women applied to more selective departments. In the Portuguese data the paradox appears in its milder, magnitude form. Within-group and population slopes share a sign, but the population exceeds all of its components. The decomposition converts the paradox from a curiosity into an accounting statement. Equation (6) says exactly how much of the population slope comes from the subgroup component and how much from the gap term.

D.2 A stylized example

A stylized numerical example illustrates the mechanism. Take three equal-size groups with widely separated but stable mean incomes, and give each group only weak within-group persistence (ρ_g between 0.23 and 0.35). Individuals churn substantially inside their own group,

while the groups stay put relative to one another. Figure 5 plots the resulting $(y_{i,t}, y_{i,t+h})$ pairs. Within each group the cloud is nearly circular and the fitted slope is shallow, yet the three clouds line up along the 45° line, so the pooled slope is steep. The population slope is $\rho = 0.90$, above *every* group slope. Equation (6) accounts for it exactly, the subgroup component $\sum_g \omega_g \rho_g$ is only 0.04 (the subgroup weights sum to just 0.14), while the between term $\omega^{GAP} r^{GAP} = 0.86$ carries almost the whole aggregate, because ω^{GAP} scales with the squared mean gap and the gap ratio $r^{GAP} \approx 1$ when the gaps are stable. Shrinking the gaps rotates the pooled line down toward the group lines; the Portuguese education partition is the mild, real-data version of the same configuration.

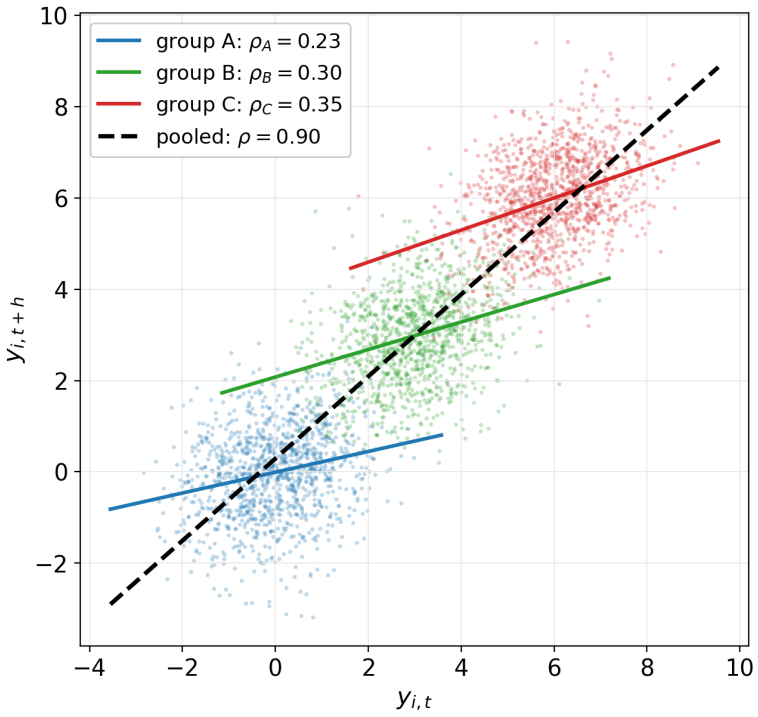


Figure 5: Simpson's paradox in a stylized panel, pooled persistence exceeds every group's

Notes: Artificial two-period panel of three equal-size groups ($N = 8{,}000$ each) with stable mean gaps ($\mu = 0, 3, 6$) and low within-group persistence. Coloured lines are within-group regressions of $y_{i,t+h}$ on $y_{i,t}$ (slopes $\rho_g = 0.23, 0.30, 0.35$); the dashed line is the pooled regression ($\rho = 0.90$). Fed to the level decomposition of equation (6): population $\rho = 0.90 = 0.04$ (subgroup) $+0.86$ (between-gap), closing to machine precision.

E Full results by partition

Table 5 reports the average and smallest cell sizes of every partition used in the paper.

Table 5: Cell sizes by partition (workers observed at t and $t+5$)

Partition	K	Mean cell size	Smallest cell	(year)
gender	2	50,512	17,312	1987
age	3	33,674	11,885	1989
education	3	33,953	1,266	1987
geography	3	33,675	14,363	1986
industry	2	50,492	20,276	1987
gender×age	6	16,837	2,458	1987
gender×education	6	16,976	253	1986
gender×geography	6	16,837	3,665	1986
gender×industry	4	25,246	6,465	1987
age×education	9	11,317	315	1986
age×geography	9	11,225	3,503	1987
age×industry	6	16,831	5,263	1987
education×geography	9	11,318	174	1986
education×industry	6	16,969	479	1994
geography×industry	6	16,831	5,183	1987
gender×age×education	18	5,659	31	1986

Notes: Mean cell size over all base years and groups, and the smallest cell with its base year, for every partition used in the paper. Cells are smallest in the 1980s, when few workers held higher education.

E.1 One-dimensional partitions

Age. The age partition is the one one-dimensional partition in which the composition-flavored channels are sizeable, precisely because the workforce aged so much over the sample (Tables 6 and 7, Figure 6). Age is a fixed attribute of the cohort entering each five-year window. A worker counts as "≥44" if they were at least 44 at t, whatever their age at $t+h$. Base-year classification is used for every partition, but the distinction between the two dates is starkest for age, since every worker mechanically ages by h years across the window. Both the compositional channel (+0.027 by 2014) and the inequality channel (+0.081) are positive. As the workforce aged, population weight and variance-share weight both shifted toward older workers, whose wages are the most persistent. They are partly offset by a mobility-*enhancing* group-gap channel

(−0.068), so the three between channels together contribute +0.040, three fifths of the +0.065 change. The within channel (+0.025 by 2014, +0.054 at the 2010 peak) is positive but smaller than under any other one-dimensional partition. This contrast is a feature of the framework, not a contradiction. As Section 3.1.1 stresses, the population change is invariant to the partition but its *attribution* is not, and the age cut is the one that surfaces the aging-composition effect the gender cut cannot see.

Table 6: Level decomposition by age bracket (rank-rank slope, $h = 5$)

Year	ρ	Subgroup	$\rho_{\leq 31}$	$\rho_{32\text{–}43}$	$\rho_{\geq 44}$	$\omega_{\leq 31}$	$\omega_{32\text{–}43}$	$\omega_{\geq 44}$	ω^{GAP}	r^{GAP}
1986	0.720	0.639	0.660	0.787	0.797	0.381	0.297	0.192	0.130	0.624
1994	0.779	0.714	0.702	0.836	0.857	0.380	0.323	0.207	0.090	0.715
2004	0.787	0.778	0.718	0.827	0.851	0.355	0.378	0.248	0.019	0.427
2011	0.810	0.796	0.711	0.838	0.865	0.252	0.416	0.311	0.021	0.618
2014	0.785	0.772	0.659	0.820	0.850	0.217	0.409	0.346	0.029	0.450

Notes: Decomposition of equation (6). "Subgroup" $= \sum_g \omega_g \rho_g$; $\omega^{GAP} = 1 - \sum_g \omega_g$ is the total gap weight and r^{GAP} the ω-weighted average gap ratio, so that $\rho = \text{Subgroup} + \omega^{GAP} r^{GAP}$ exactly.

Table 7: Four-channel change decomposition (age partition)

Year t	$\Delta\rho$	Within	Compositional	Inequality	Group-gap
1991	+0.0417	+0.0320	-0.0012	+0.0051	+0.0057
1994	+0.0592	+0.0433	-0.0006	+0.0328	-0.0163
2004	+0.0671	+0.0466	+0.0029	+0.0904	-0.0728
2011	+0.0901	+0.0509	+0.0180	+0.0889	-0.0678
2014	+0.0652	+0.0252	+0.0273	+0.0808	-0.0680

Notes: Decomposition of equation (9), age partition, base year $t_0 = 1986$, horizon $h = 5$, Shapley attribution.

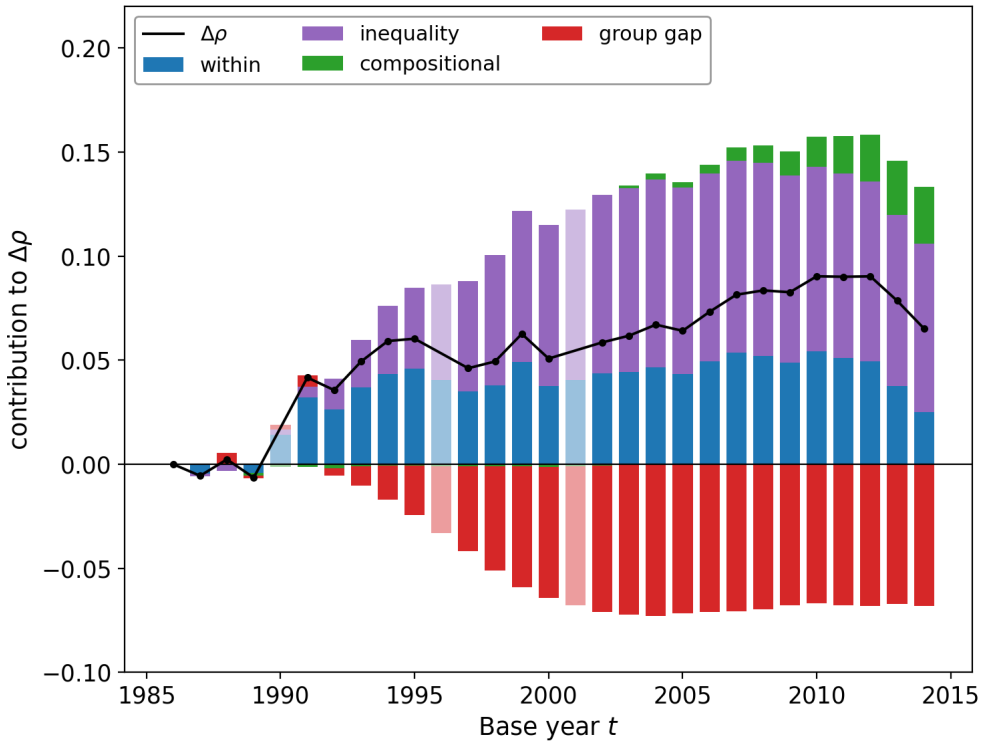


Figure 6: Change decomposition relative to 1986 (age)

Notes: Cumulative contribution of each channel of equation (9) to $\Delta\rho$ relative to the 1986 base year, age partition, Shapley attribution; the base year 1986 is fixed at zero. Base years 1990, 1996 and 2001 have no five-year pair and are linearly interpolated (lighter bars).

The higher-persistence group here is older workers ($\rho_{\geq 44}$ around 0.80–0.87), the lower-persistence group the young ($\rho_{\leq 31}$ around 0.66–0.72). The blue "within" bars are positive but the smallest of any one-dimensional partition, and they shrink by 2014, because the youngest tercile never became stickier whereas the middle and older terciles did. This is evidenced by $\rho_{\leq 31}$ ending at 0.66, essentially its 1986 value, along with the flat young line against the rising middle and older lines in the persistence (ρ_t) panel of Figure 22; the within rise is carried by the 32–43 and ≥ 44 groups. The defining feature of the age cut is that the green "compositional" and purple "inequality" bars are *both* positive. As the workforce aged, both population weight and variance-share weight shifted from the low-persistence young toward the high-persistence old, the young tercile's population share fell from 0.41 to 0.27 and its variance share from 0.93 to 0.81, with mirror-image gains among older workers, so that both the composition (n_t) and variance-share (v_t) panels of Figure 22 tilt away from the young toward the old. The

compositional bar is near zero early and grows only after about 2004, when aging accelerated, whereas the inequality bar is sizeable from the mid-1990s and peaks around 2004. Finally, the red "group-gap" bars are tiny and positive in 1991 and then strongly negative (mobility-*enhancing*). The between-age-group rank gaps carry less persistence than in 1986, since as each fixed cohort ages across its five-year window the young partly catch up and the relative-rank gaps compress, echoing how the age profile of mean wages in the μ_t panel of Figure 22 bunches together across a window; this offsets most of the aging-driven composition and inequality rise. The net change matches the population change and coincides with the gender partition, since age and gender are recorded for every worker and the aggregate change is partition-invariant.

Education. The education partition (Tables 8 and 9, Figure 7) reveals two large, offsetting between-group forces that the gender partition cannot see. The compositional channel is strongly mobility-enhancing (−0.087 by 2014). Educational upskilling moved population weight, and with it the variance-weighted persistence of the basic-education group, out of the most persistent group, exactly the mechanism of Observation 2; the inequality channel is small (+0.020). Against this stands the group-gap channel (+0.074). As the educated groups expanded, the large and persistent wage gaps between education groups came to carry more weight in the population. The within channel (+0.060) tells the same story as before, churning declined inside every education group. The net effect is a decline in mobility in which within-group persistence does nine tenths of the work, the three between channels netting to +0.007.

Table 8: Level decomposition by education (rank-rank slope, $h = 5$)

Year	ρ	Subgroup	ρ_{Basic}	$\rho_{Secondary}$	ρ_{Higher}	ω_{Basic}	$\omega_{Secondary}$	ω_{Higher}	ω^{GAP}	r^{GAP}
1986	0.718	0.576	0.688	0.594	0.640	0.721	0.128	0.005	0.145	0.979
1994	0.771	0.604	0.715	0.732	0.589	0.598	0.239	0.005	0.159	1.046
2004	0.788	0.574	0.722	0.727	0.661	0.378	0.378	0.041	0.204	1.047
2011	0.810	0.582	0.743	0.756	0.763	0.238	0.446	0.089	0.227	1.004
2014	0.785	0.569	0.740	0.710	0.703	0.200	0.471	0.122	0.206	1.050

Notes: Decomposition of equation (6) with $K = 3$ education groups (basic ≤6 years; secondary 7–12; higher >12); the gap term sums over the three group pairs (Appendix A). The last two columns report the total gap weight $\omega^{GAP} = 1 - \sum_g \omega_g$ and the ω-weighted average gap ratio r^{GAP}, so that $\rho = \text{Subgroup} + \omega^{GAP} r^{GAP}$ exactly. Education is unrecorded in 1995, 1997, and 1998, so those years drop out of the education partition.

Table 9: Four-channel change decomposition (education partition)

Year t	$\Delta\rho$	Within	Compositional	Inequality	Group-gap
1991	+0.0434	+0.0259	-0.0150	-0.0224	+0.0549
1994	+0.0529	+0.0421	-0.0091	-0.0045	+0.0244
2004	+0.0694	+0.0513	-0.0507	-0.0020	+0.0708
2011	+0.0917	+0.0757	-0.0766	+0.0073	+0.0853
2014	+0.0667	+0.0598	-0.0869	+0.0200	+0.0738

Notes: Decomposition of equation (9), education partition ($K = 3$), base year $t_0 = 1986$, Shapley attribution. Education is unrecorded in 1995, 1997, and 1998, so those years drop out of the education partition.

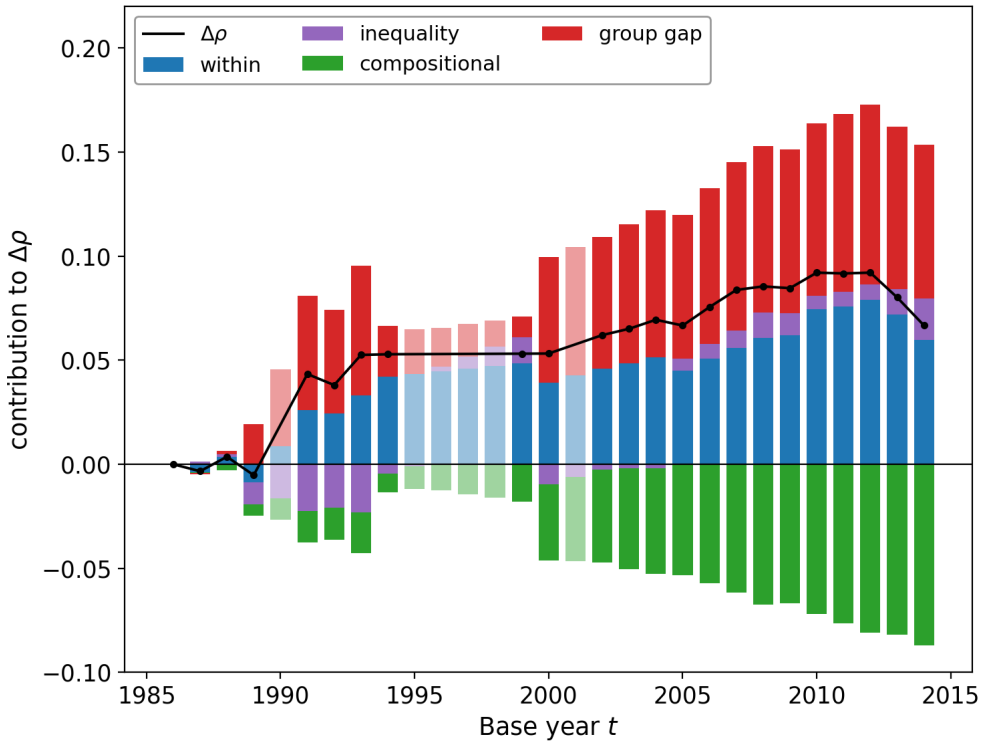


Figure 7: Change decomposition relative to 1986 (education)

Notes: Cumulative contribution of each channel of equation (9) to $\Delta\rho$ relative to the 1986 base year, education partition, Shapley attribution. Base years 1990, 1995–1998 and 2001 (no survey, or education unrecorded) are linearly interpolated and drawn lighter.

The education partition uses a slightly different sample, workers with recorded base-year education, excluding the 1995 and 1997–98 waves, so its net $\Delta\rho$ is close to, but not identical with, the gender and age series. Historically the high-persistence group is basic education ($\rho_{Basic} = 0.69$ in 1986), which was also by far the largest (79% of workers) and held most of the wage variance (variance share 0.91). The blue "within" bars are positive and growing. Every education group became stickier (ρ_{Basic} 0.69→0.74, $\rho_{Secondary}$ 0.59→0.71, ρ_{Higher} 0.64→0.70), as all three lines rise in the persistence (ρ_g) panel of Figure 27. The green "compositional" bars are *negative* and growing (mobility-*enhancing*). The great upskilling moved population weight out of the high-persistence basic group, whose share fell from 79% to 26%, the collapse of the basic-education line in the composition (n_g) panel of Figure 27. The purple "inequality" bars are small, because the basic group's variance share fell only from 0.91 to 0.77 while the expanding groups' shares rose, movements that largely offset once weighted by pop-

ulation share and persistence. How the re-weighting is split between the two channels depends on the attribution rule, since shares and variance shares moved together (Appendix F); their sum remains negative (−0.067 under Shapley). Finally, the red "group-gap" bars are large and *positive* (mobility-reducing). As the secondary and higher groups grew from the margins (higher education went from 2% to 21% of workers), the large and persistent wage gaps *between* education groups came to carry more weight in the aggregate, the wide, stable log-wage gaps in the mean-gap panel of Figure 32. The re-weighting and the gap channel are thus large and offsetting, the first mobility-enhancing and the second mobility-reducing, which leaves within-group persistence as the net driver under this partition.

Region. The three broad regions (Tables 10 and 11, Figure 8) attribute the whole change to within-region persistence. The within channel is +0.068 of +0.065 by 2014, the inequality channel +0.040 and the group-gap channel −0.039 offset each other, and the compositional channel is negligible (−0.004). Regional wage gaps are small to begin with and narrowed over the period, so the gap term of the level decomposition falls from 0.097 to 0.058.

Table 10: Level decomposition by region (rank-rank slope, $h = 5$)

Year	ρ	Subgroup	ρ_{Lisbon}	ρ_{North}	$\rho_{Centre-South}$	ω_{Lisbon}	ω_{North}	$\omega_{Centre-South}$	ω^{GAP}	r^{GAP}
1986	0.720	0.622	0.703	0.716	0.672	0.277	0.366	0.246	0.110	0.881
1994	0.779	0.682	0.788	0.752	0.739	0.256	0.359	0.285	0.100	0.963
2004	0.787	0.710	0.783	0.800	0.724	0.269	0.337	0.318	0.076	1.000
2011	0.810	0.744	0.831	0.808	0.767	0.271	0.354	0.304	0.072	0.918
2014	0.785	0.726	0.769	0.802	0.738	0.292	0.350	0.299	0.058	0.999

Notes: Decomposition of equation (6). "Subgroup" $= \sum_g \omega_g \rho_g$; $\omega^{GAP} = 1 - \sum_g \omega_g$ is the total gap weight and r^{GAP} the ω-weighted average gap ratio, so that $\rho = \text{Subgroup} + \omega^{GAP} r^{GAP}$ exactly.

Table 11: Four-channel change decomposition (region partition)

Year t	$\Delta\rho$	Within	Compositional	Inequality	Group-gap
1991	+0.0416	+0.0337	+0.0016	+0.0073	-0.0009
1994	+0.0592	+0.0538	+0.0008	+0.0051	-0.0005
2004	+0.0671	+0.0664	-0.0033	+0.0247	-0.0208
2011	+0.0901	+0.0948	-0.0034	+0.0302	-0.0315
2014	+0.0652	+0.0681	-0.0035	+0.0395	-0.0388

Notes: Decomposition of equation (9), region partition, base year $t_0 = 1986$, horizon $h = 5$, Shapley attribution.

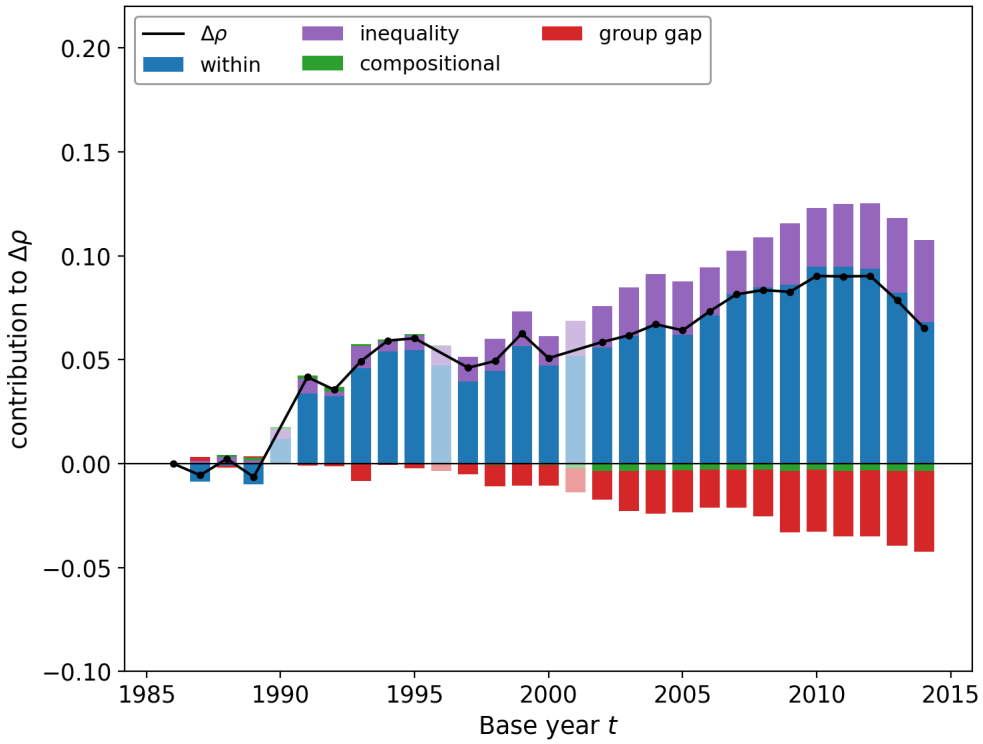


Figure 8: Four-channel change decomposition relative to 1986, geography partition ($h = 5$)
Notes: As in Figure 6, for the broad-geography partition Lisbon / North / Centre–South (NUTS-2 regions). The base year 1986 is fixed at zero. Base years without a five-year pair are interpolated (lighter bars).

Industry. The goods–services partition (Tables 12 and 13, Figure 9) tells the same story: within +0.071, inequality +0.036, group-gap −0.045, compositional +0.003. Although employment moved massively into services, the two sectors are close in persistence and in mean wages, so the shift carries almost no compositional weight.

Table 12: Level decomposition by broad industry (rank-rank slope, $h = 5$)

Year	ρ	Subgroup	ρ_{Goods}	$\rho_{Services}$	ω_{Goods}	$\omega_{Services}$	ω^{GAP}	r^{GAP}
1986	0.720	0.663	0.708	0.714	0.583	0.351	0.067	0.849
1994	0.779	0.728	0.740	0.786	0.477	0.477	0.046	1.101
2004	0.787	0.760	0.772	0.785	0.388	0.587	0.025	1.055
2011	0.810	0.792	0.800	0.812	0.331	0.650	0.019	0.897
2014	0.785	0.773	0.801	0.772	0.321	0.669	0.010	1.123

Notes: Decomposition of equation (6). "Subgroup" $= \sum_g \omega_g \rho_g$; $\omega^{GAP} = 1 - \sum_g \omega_g$ is the total gap weight and r^{GAP} the ω-weighted average gap ratio, so that $\rho =$ Subgroup $+ \omega^{GAP} r^{GAP}$ exactly.

Table 13: Four-channel change decomposition (broad-industry partition)

Year t	$\Delta\rho$	Within	Compositional	Inequality	Group-gap
1991	+0.0415	+0.0297	+0.0011	+0.0030	+0.0077
1994	+0.0590	+0.0466	+0.0043	+0.0141	-0.0060
2004	+0.0668	+0.0635	+0.0068	+0.0262	-0.0297
2011	+0.0899	+0.0904	+0.0056	+0.0333	-0.0394
2014	+0.0650	+0.0708	+0.0033	+0.0359	-0.0450

Notes: Decomposition of equation (9), broad-industry partition, base year $t_0 = 1986$, horizon $h = 5$, Shapley attribution.

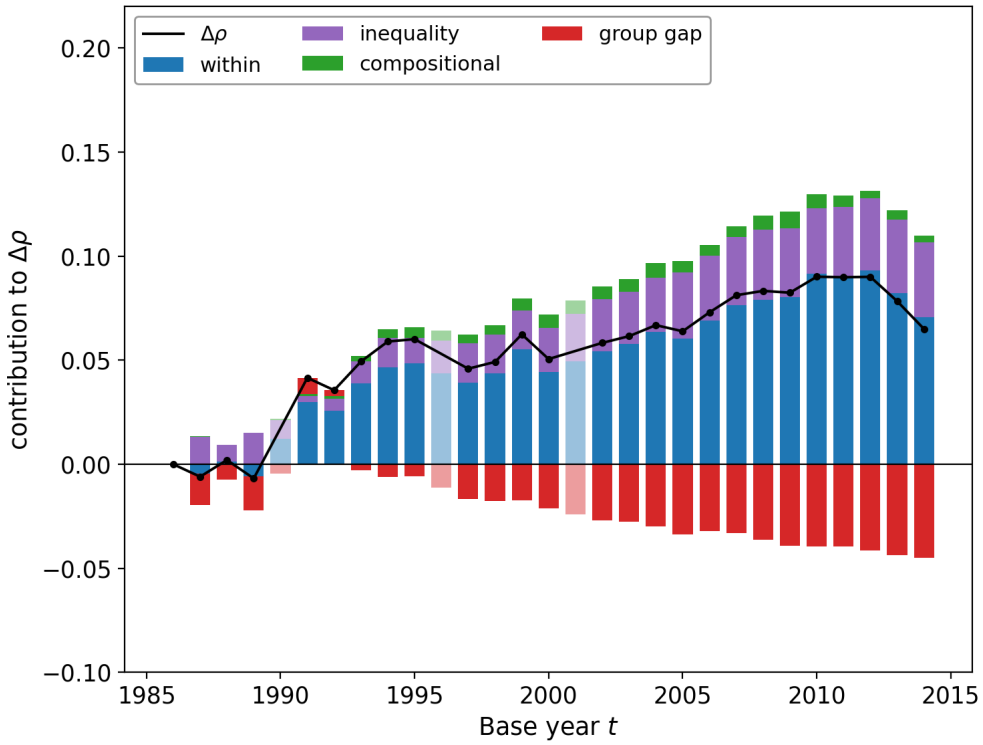


Figure 9: Four-channel change decomposition relative to 1986, broad-industry partition ($h = 5$)

Notes: As in Figure 6, for the broad-industry partition Goods (sections A–F) vs. Services (sections G–U). The base year 1986 is fixed at zero. Base years without a five-year pair are interpolated (lighter bars).

E.2 Joint partitions

Gender ×age ×education. The one-dimensional subgroup decompositions attribute a large part of the rise in population wage persistence to the within channel. This need not be the end of the story. The within-gender rise, for instance, could itself partly reflect mean-gap changes along *other* dimensions. If the education wage premium widens within gender, workers of different education sort further apart inside each gender pool, and the one-dimensional gender decomposition attributes that widening to within-gender persistence rather than to a gap between groups. To absorb these cross-dimensional interactions, I cut the population finely and let the unit of churning be the joint gender×age×education *cell*. Young, higher-educated women churn among themselves, older, basic-educated men among themselves, and so on.

Whether this fine cell is the right level of analysis is a question I return to in Section 5.

Crossing the three age terciles with gender and the three education groups splits the population into $K = 18$ cells, each worker assigned by age, gender and education at the base year t. Because the cut uses education, it follows the education-complete flow (population percentile ranks over all workers; a worker enters at a base year when base-year education is recorded), so the population change $\Delta\rho$ coincides exactly with the education partition's (Table 9) rather than with gender's or age's. The fine cut changes the attribution sharply (Table 14, Figure 10). The within channel is +0.014 at the 2011 peak and −0.011 by 2014. Once workers are grouped this finely, no within-cell slowdown of churning is left to move the aggregate. The inequality channel carries most of the change (+0.054 in 2011, +0.070 in 2014), as variance-share weight moved toward the more persistent cells, with the compositional channel close to zero and the group-gap channel positive throughout (+0.073 in 1991, +0.028 in 2011, +0.010 in 2014), the persistent wage gaps between cells of different education. The smallest cells of this partition are thin in the 1980s (Table 5), which is why the paper's punchline rests on the nine-cell age×education partition of Section 4.2; the eighteen-cell partition confirms it.

Table 14: Four-channel change decomposition (gender×age×education partition, $K = 18$)

Year t	$\Delta\rho$	Within	Compositional	Inequality	Group-gap
1991	+0.0434	+0.0076	-0.0088	-0.0283	+0.0728
1994	+0.0529	+0.0141	+0.0005	+0.0063	+0.0319
2004	+0.0694	+0.0143	-0.0042	+0.0333	+0.0261
2011	+0.0917	+0.0139	-0.0040	+0.0535	+0.0282
2014	+0.0667	-0.0106	-0.0019	+0.0695	+0.0097

Notes: Decomposition of equation (9) for the joint gender×age×education partition (eighteen cells: the three age terciles crossed with gender and with basic, secondary and higher education), from base year $t_0 = 1986$ to end year t_1, horizon $h = 5$, Shapley attribution. Ranks are population percentile ranks over all workers; a worker enters when base-year education is recorded, so the $\Delta\rho$ column matches the education partition (Table 9). The four channels sum to $\Delta\rho$ exactly, up to rounding. Years 1990 and 2001 are absent (no survey), and education is unrecorded in 1995, 1997, and 1998, so those base years drop out.

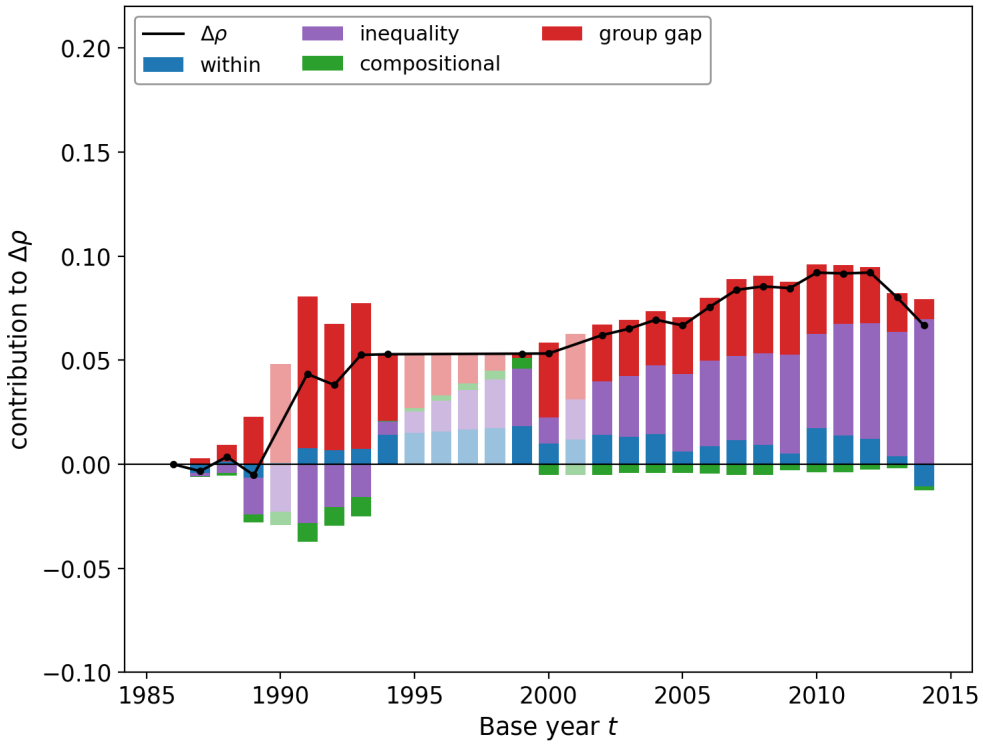


Figure 10: Change decomposition relative to 1986 (gender×age×education)

Notes: Cumulative contribution of each channel of equation (9) to $\Delta\rho$ relative to the 1986 base year, joint gender×age×education partition ($K = 18$), Shapley attribution; the base year 1986 is fixed at zero. The line is the population change; the bars are the four channels. Base years without a five-year pair or without recorded education are interpolated (lighter bars).

Two-way partitions. Table 15 reports the decomposition for every one-dimensional partition and every pair of characteristics at 2014, on the same bracketing throughout. All recover essentially the same population change (+0.065 to +0.067, the education-based cuts using the education-complete sample), but they attribute it very differently. Under gender×education (Figure 11) the group-gap channel is positive (+0.046) and the compositional channel negative (−0.038), the education pattern of Figure 7 carried through, and the within channel keeps +0.046. Under gender×age (Figure 12) the group-gap channel reverses sign, to −0.087, offset by compositional (+0.058) and inequality (+0.067) terms. Age and gender mean gaps compressed over the period while education gaps did not, so whether the between channels raise or lower measured persistence depends on which characteristics the partition uses. Every pair that involves age halves the within channel (+0.028 to +0.031); only age×education removes

it (+0.004), as discussed in Section 4.2. Table 16 splits the between channels of that partition into their age- and education-driven parts.

Table 15: Change decomposition 1986 → 2014, every one-dimensional and pairwise partition

Partition	K	$\Delta\rho$	Within	Compositional	Inequality	Group-gap	Within share
Gender	2	+0.065	+0.065	+0.024	+0.020	-0.044	1.00
Age	3	+0.065	+0.025	+0.027	+0.081	-0.068	0.39
Education	3	+0.067	+0.060	-0.087	+0.020	+0.074	0.90
Region	3	+0.065	+0.068	-0.004	+0.039	-0.039	1.04
Industry	2	+0.065	+0.071	+0.003	+0.036	-0.045	1.09
Gender (education-complete)	2	+0.067	+0.066	+0.024	+0.020	-0.044	1.00
Gender×Age	6	+0.065	+0.028	+0.058	+0.067	-0.087	0.42
Gender×Education	6	+0.067	+0.046	-0.038	+0.012	+0.046	0.70
Gender×Region	6	+0.065	+0.068	+0.016	+0.054	-0.073	1.05
Gender×Industry	4	+0.065	+0.066	+0.043	+0.038	-0.083	1.01
Age×Education	9	+0.067	+0.004	-0.044	+0.094	+0.013	0.06
Age×Region	9	+0.065	+0.030	+0.027	+0.092	-0.084	0.46
Age×Industry	6	+0.065	+0.031	+0.032	+0.102	-0.100	0.48
Education×Region	9	+0.067	+0.062	-0.072	+0.037	+0.040	0.93
Education×Industry	6	+0.066	+0.062	-0.077	+0.039	+0.043	0.93
Region×Industry	6	+0.065	+0.069	+0.019	+0.047	-0.070	1.06
Gender×Age×Education	18	+0.067	-0.011	-0.002	+0.070	+0.010	-0.16

Notes: Four-channel decomposition of equation (9) from 1986 to 2014 ($h = 5$), Shapley attribution, for the five characteristics taken one at a time, all ten pairs, and the three-way gender×age×education cut. "Within share" is the within channel divided by $\Delta\rho$. Partitions that use education are computed on the education-complete sample, whose change is +0.067 rather than +0.065.

Table 16: Between-group change of the age×education partition, by dimension

Channel	B(age)	B(edu)	B(age × edu)	ϕ_{age}	ϕ_{edu}	Interaction
1986 → 2011: $\Delta\rho = +0.092$; within: age +0.052, education +0.076, age×education +0.024						
Compositional	+0.018	-0.077	-0.042	+0.026	-0.068	+0.016
Inequality	+0.089	+0.007	+0.077	+0.079	-0.003	-0.020
Group-gap	-0.068	+0.085	+0.033	-0.060	+0.093	+0.015
Between (total)	+0.039	+0.016	+0.067	+0.045	+0.022	+0.012
gap by pair type	age pairs -0.047, education pairs +0.033, cross pairs +0.046					
1986 → 2014: $\Delta\rho = +0.067$; within: age +0.027, education +0.060, age×education +0.004						
Compositional	+0.027	-0.087	-0.044	+0.035	-0.079	+0.016
Inequality	+0.081	+0.020	+0.094	+0.077	+0.016	-0.008
Group-gap	-0.068	+0.074	+0.013	-0.064	+0.077	+0.007
Between (total)	+0.040	+0.007	+0.063	+0.048	+0.015	+0.016
gap by pair type	age pairs -0.050, education pairs +0.027, cross pairs +0.036					

Notes: On the education-complete sample, $B(S)$ is the contribution of a channel under partition S (age alone, education alone, or the two crossed), Shapley attribution. $\phi_{\text{age}} = \frac{1}{2}[B(\text{age}) + B(\text{age} \times \text{edu}) - B(\text{edu})]$ and ϕ_{edu} symmetrically; they sum to $B(\text{age} \times \text{edu})$. The interaction $B(\text{age} \times \text{edu}) - B(\text{age}) - B(\text{edu})$ is split equally between the two. The last row of each block splits the joint group-gap channel exactly by pair type: pairs of cells that differ in age only, in education only, or in both.

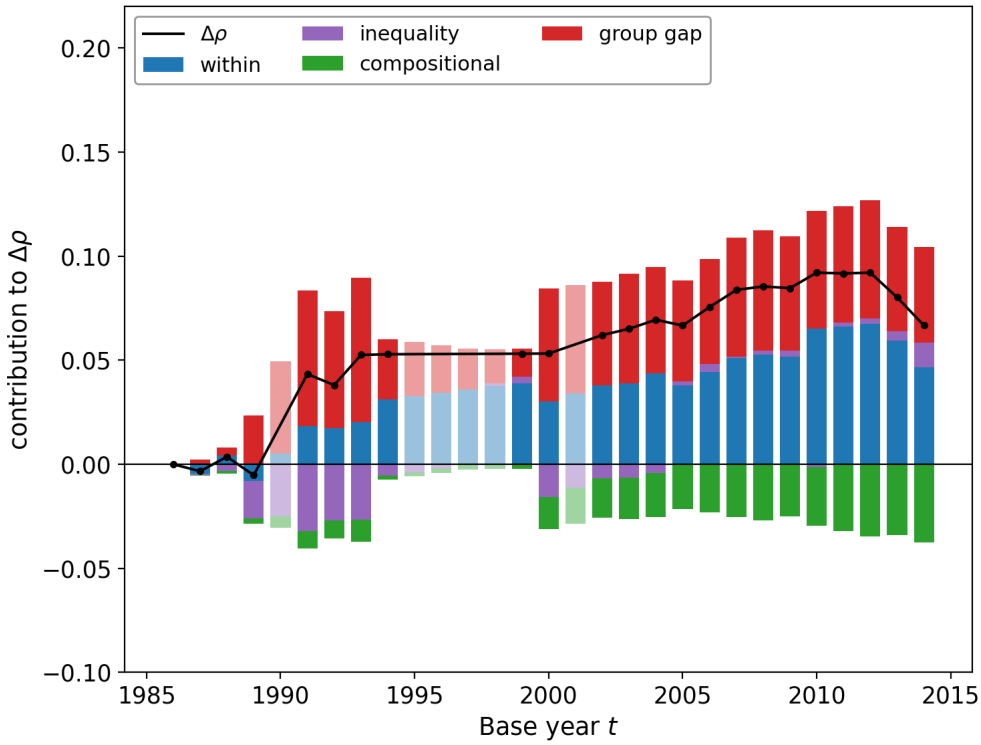


Figure 11: Change decomposition relative to 1986 (gender×education)

Notes: As in Figure 4, for the gender×education partition ($K = 6$). The group-gap channel is positive throughout, as in the one-dimensional education partition of Figure 7, while the compositional channel turns increasingly negative as upskilling moves weight away from the basic-education cells. Base years without a five-year pair or without recorded education are interpolated (lighter bars).

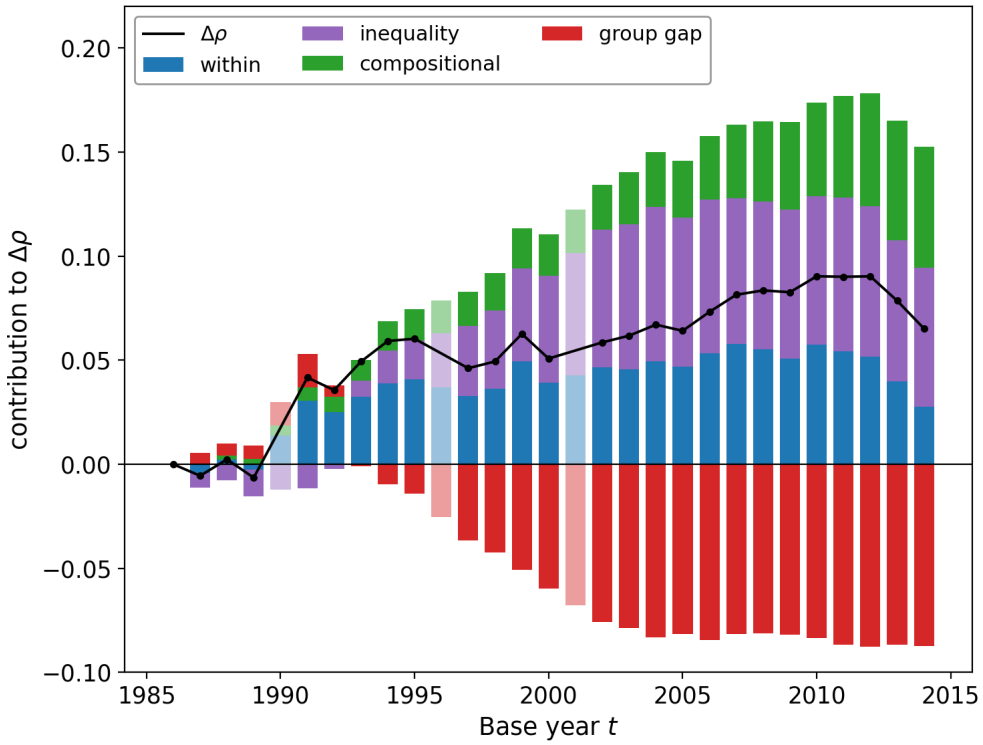


Figure 12: Change decomposition relative to 1986 (age×gender)

Notes: As in Figure 4, for the age×gender partition ($K = 6$), the three age terciles crossed with gender. Education is not used to define the cells, so no base years are lost to the unrecorded-education waves. In contrast to the two partitions that include education, the group-gap channel here is strongly *negative*, age and gender mean gaps compressed over the period, and is offset by large positive compositional and inequality channels. The sign of the between channels therefore depends on which characteristics the partition uses, which is the point of Observation 1.

E.3 The correlation of log wages and permanent income

Correlation of log wages. Table 17 sets the slope and the correlation decompositions side by side for each partition, and Figures 13–15 plot the correlation versions. The qualitative verdict is the same under both measures. The within channel carries the whole change under gender, region and industry, about half of it under age, and vanishes under age×education (−0.001 of +0.044); the compression of the gender gap and of the age gaps is mobility-enhancing under both.

Where the two measures differ is in *attribution*. The correlation weights each group by its share of *log-wage dispersion* ($\sigma_{g,t}$), whereas the slope weights it by its share of rank variance.

Log-wage dispersion moved substantially across groups over the sample, as the σ_t panels of Figures 21–24 show. The consequences differ by partition. Under *gender*, the correlation's dispersion channel is larger than the slope's inequality channel (+0.048 versus +0.020 by 2014) and its compositional channel flips sign (−0.018 against +0.024). Because women are *less* dispersed in log wages than men (the σ_t panel of Figure 21), feminization now shifts weight toward a lower-dispersion group. Under *age*, the slope loads the aging effect onto a large inequality channel (+0.081), but the log-wage dispersion-share shift toward older workers is far smaller than the rank-variance shift (the dispersion-share panel of Figure 31); the correlation's dispersion channel therefore is small (+0.015), and the aging effect loads onto the compositional channel instead (+0.064 versus +0.027 for the slope).

Education shows the sharpest divergence. The correlation splits the change into two enormous, offsetting re-weighting channels, compositional positive (+0.16 by 2014) and dispersion negative (−0.17), and leaves the within channel with +0.024 of the +0.044 change, where the slope assigns +0.060 of +0.067 to within and offsets a compositional channel of −0.087 with a group-gap channel of +0.074. The reason is that upskilling dramatically reallocated log-wage dispersion across education groups. The small but highly dispersed higher-education group expanded (its own correlation rising over the sample) while the large basic group shrank, a reallocation visible in the population-share and dispersion-share panels of Figure 32. The direction of the change is the same across measures; the log-wage correlation attributes the rise to the changing composition and dispersion of wage inequality, while the rank-rank slope attributes it to within-cell churn and education gaps.

Table 17: Rank-rank slope and log-wage correlation decompositions, 1986 → 2014

Partition	Measure	Δ	Within	Compositional	Inequality / dispersion	Group-gap
Gender						
	rank-rank slope ρ	+0.065	+0.065	+0.024	+0.020	-0.044
	log-wage correlation r	+0.042	+0.046	-0.018	+0.048	-0.035
Age						
	rank-rank slope ρ	+0.065	+0.025	+0.027	+0.081	-0.068
	log-wage correlation r	+0.042	+0.022	+0.064	+0.015	-0.060
Education						
	rank-rank slope ρ	+0.067	+0.060	-0.087	+0.020	+0.074
	log-wage correlation r	+0.044	+0.024	+0.158	-0.168	+0.030
Region						
	rank-rank slope ρ	+0.065	+0.068	-0.004	+0.039	-0.039
	log-wage correlation r	+0.042	+0.050	-0.017	+0.045	-0.036
Industry						
	rank-rank slope ρ	+0.065	+0.071	+0.003	+0.036	-0.045
	log-wage correlation r	+0.041	+0.046	+0.068	-0.024	-0.049
Age×education						
	rank-rank slope ρ	+0.067	+0.004	-0.044	+0.094	+0.013
	log-wage correlation r	+0.044	-0.001	+0.177	-0.128	-0.004

Notes: Four-channel decompositions of the rank-rank slope (equation (9)) and of the correlation of log wages (equation (17)) from 1986 to 2014, $h = 5$, Shapley attribution. For the correlation the inequality channel is the dispersion channel.

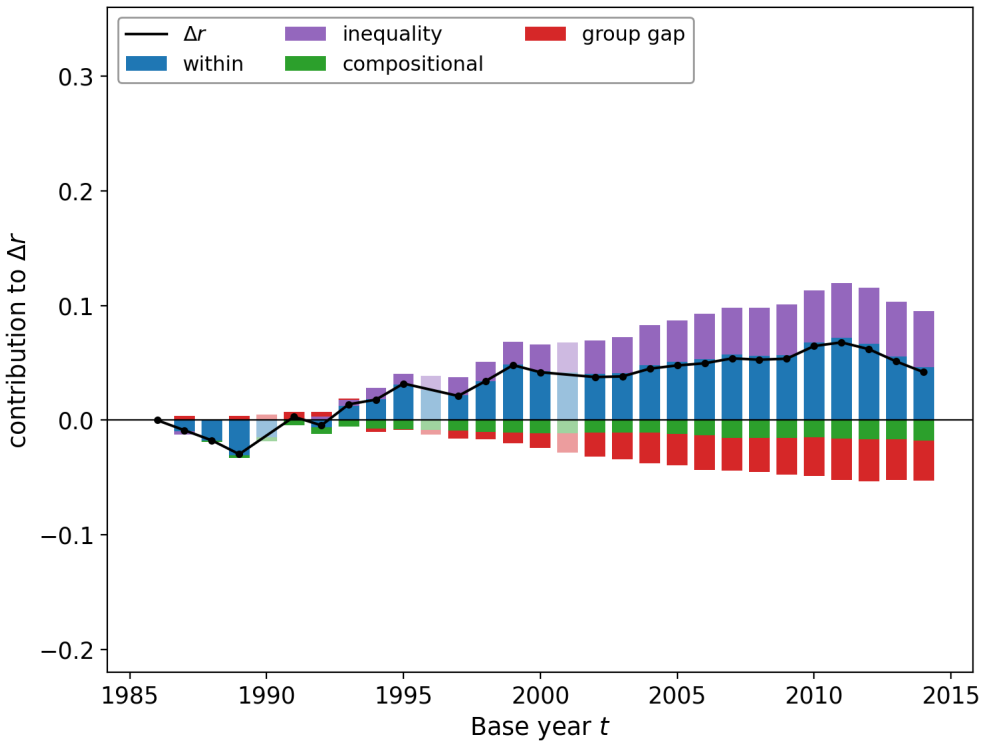


Figure 13: Change decomposition of the log-wage correlation, gender

Notes: Shapley attribution. Base years without a five-year pair are interpolated (lighter bars).

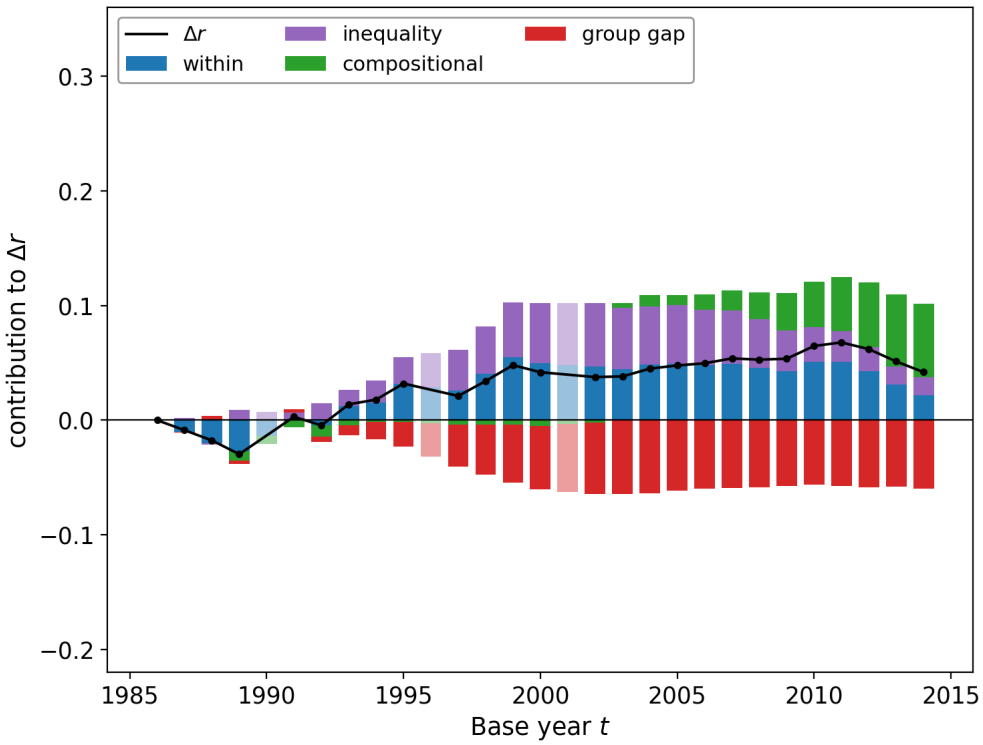


Figure 14: Change decomposition of the log-wage correlation, age

Notes: Shapley attribution. Base years without a five-year pair are interpolated (lighter bars).

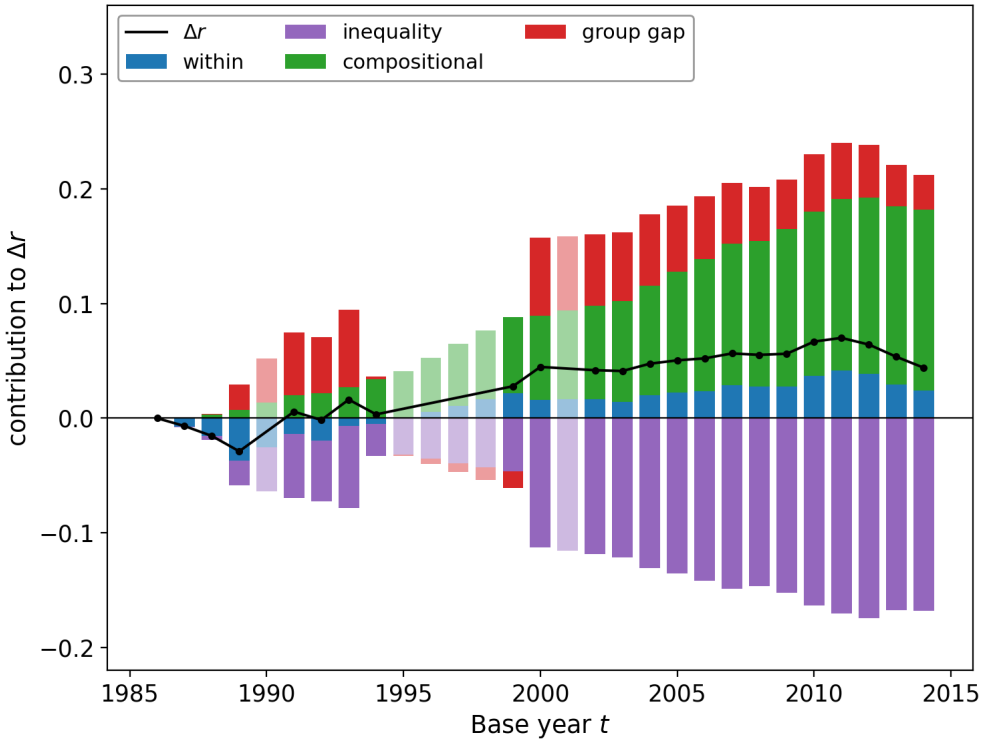


Figure 15: Change decomposition of the log-wage correlation, education

Notes: Four-channel change decomposition of the Pearson correlation of the log real wage between t and $t+5$, relative to 1986 (fixed at zero), Shapley attribution; the inequality channel is here driven by log-wage dispersion shares rather than rank-variance shares (Appendix B). Group membership is taken at the base year t. Base years without a five-year pair or without recorded education are interpolated (lighter bars).

Permanent income. Removing transitory wage fluctuations does not overturn the decline. Figure 16 repeats the five-year rank-rank exercise for "permanent income", the centred three-year mean of the log real wage, and the hump is, if anything, more pronounced. Purging transitory variation raises measured persistence throughout while leaving the trend intact. Figures 17–19 repeat the four-channel *slope* decomposition for permanent income, re-ranked within year. Because a worker's 1986 permanent income would require 1985 wages, which the panel lacks, permanent income is defined only from 1987, and the series is anchored there rather than at 1986. The decline is if anything sharper than for annual income (the net rise reaches about +0.082 by 2013), and the within channel again carries it under gender (+0.079) and education (+0.079), and two thirds of it under age (+0.054).

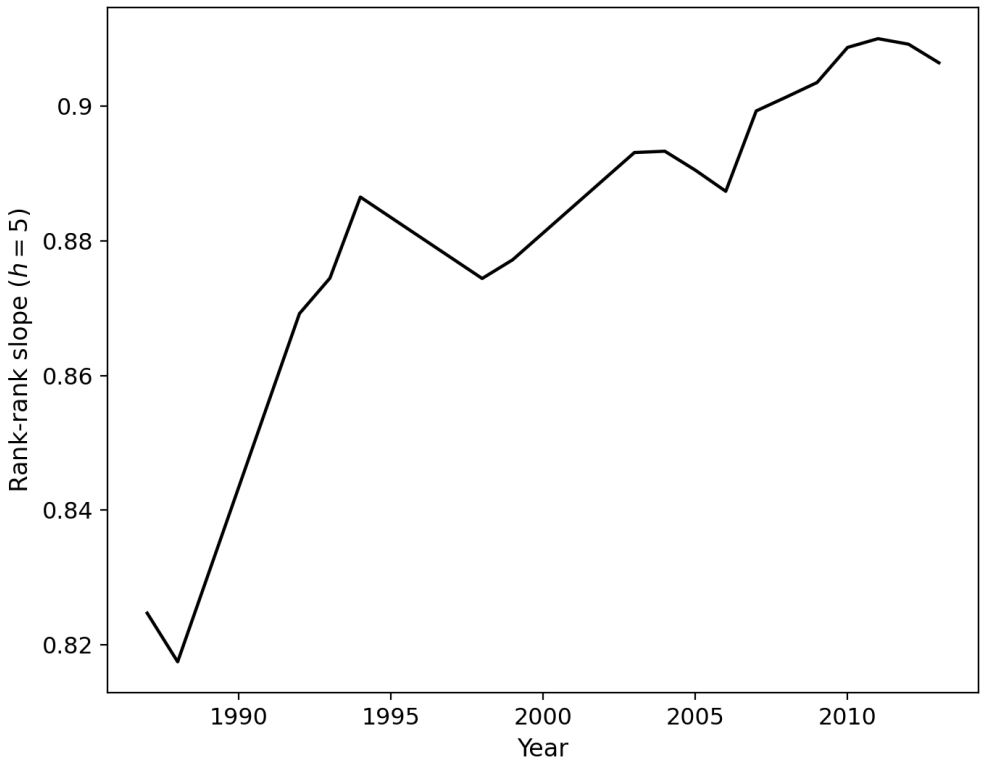


Figure 16: Rank-rank slope, permanent income (population), $h = 5$

Notes: Permanent income is the centred three-year arithmetic mean of the log real wage; ranks are computed within year over the permanent-income distribution. Slope of the regression of the rank at $t+5$ on the rank at t. Permanent income requires three consecutive annual wages, so it is unavailable in 1986 and around the missing 1990 and 2001 surveys; those base years are interpolated.

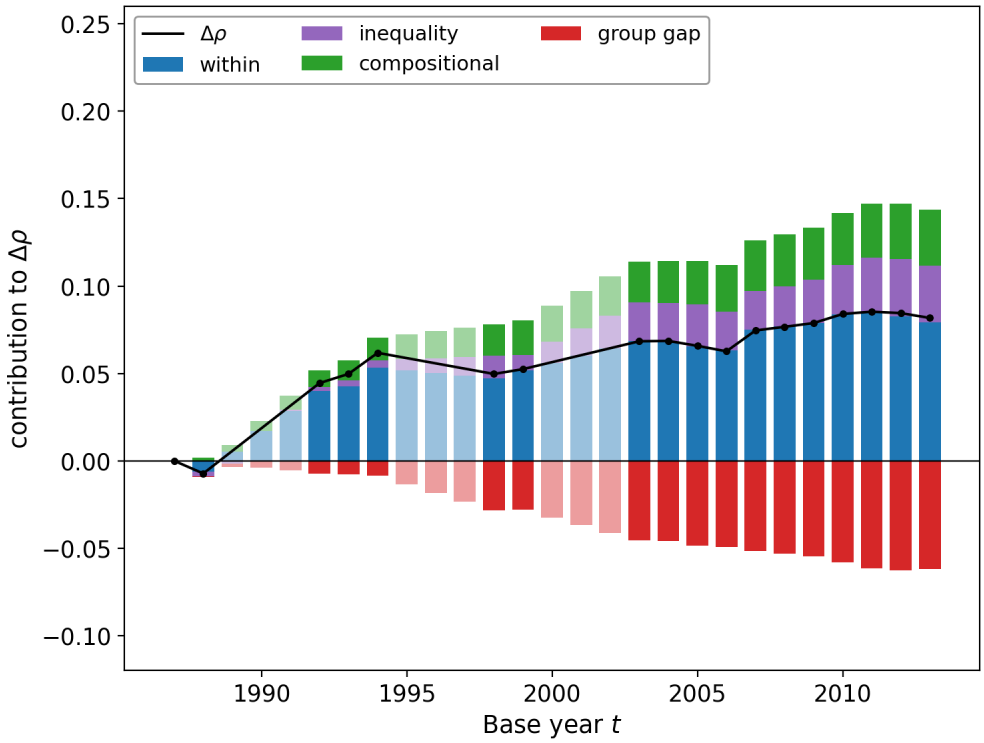


Figure 17: Change decomposition of permanent income, gender

Notes: Shapley attribution, relative to the 1987 base year. Permanent income requires three consecutive annual wages, so it is unavailable in 1986 and around the missing 1990 and 2001 surveys (interpolated, lighter bars).

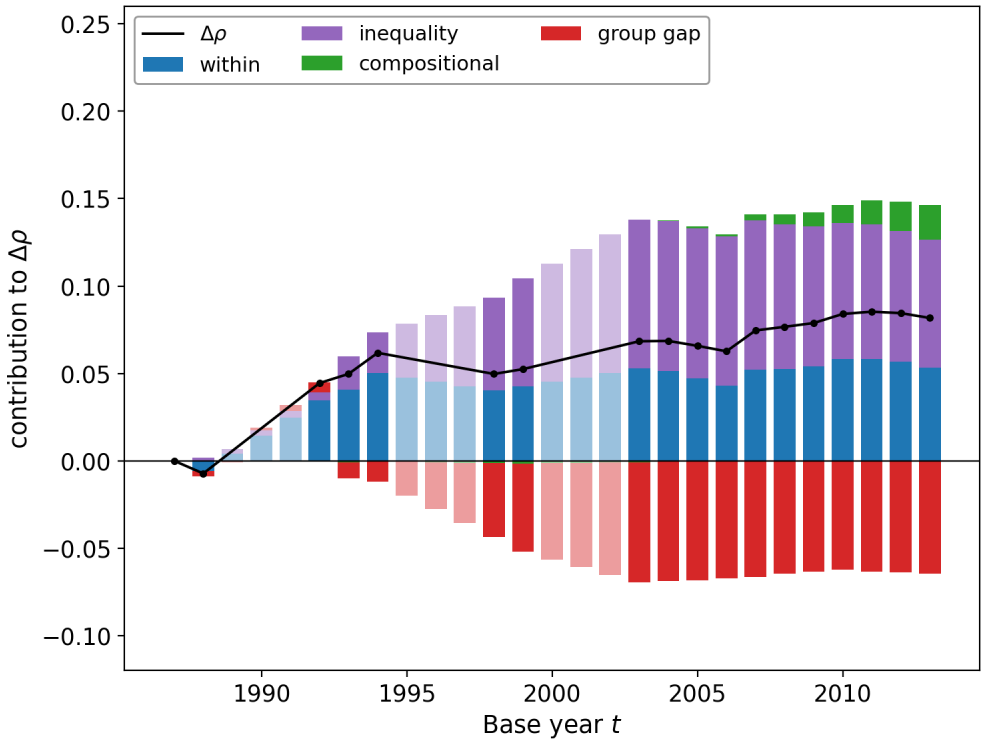


Figure 18: Change decomposition of permanent income, age

Notes: Shapley attribution, relative to the 1987 base year. Permanent income requires three consecutive annual wages, so it is unavailable in 1986 and around the missing 1990 and 2001 surveys (interpolated, lighter bars).

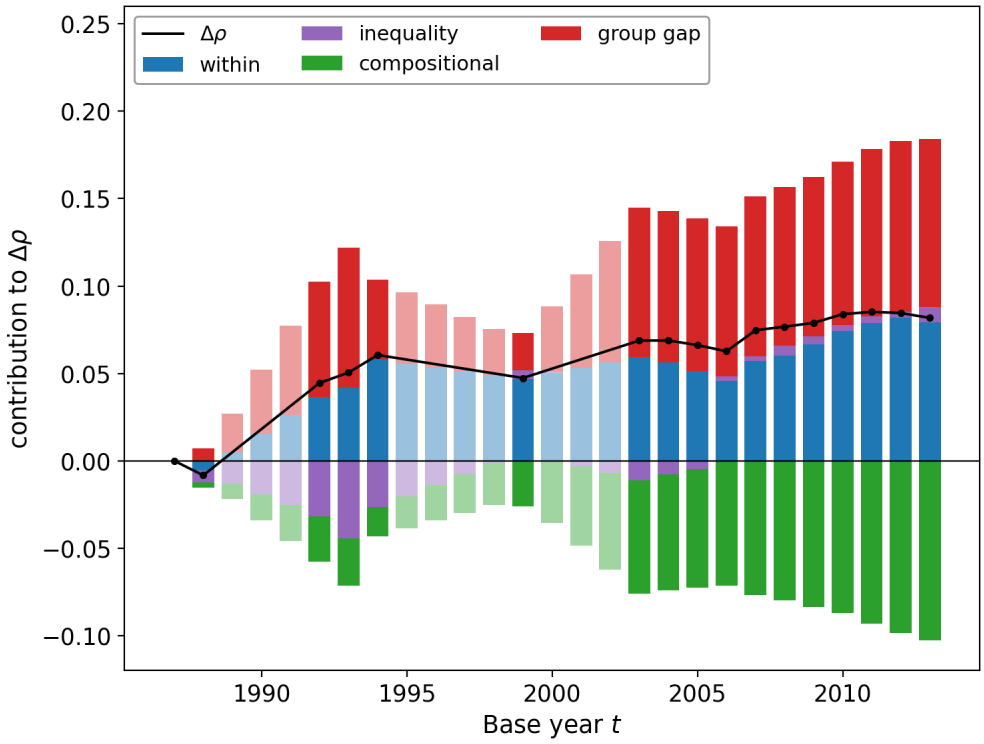


Figure 19: Change decomposition of permanent income, education

Notes: Four-channel change decomposition of the rank-rank slope of permanent income (centred three-year mean of the log real wage, re-ranked within year), relative to the 1987 base year (fixed at zero), Shapley attribution. Group membership is taken at the base year t. Permanent income needs three consecutive annual wages (unavailable in 1986 and around the 1990 and 2001 gaps); education is also unrecorded in 1995, 1997, and 1998; those base years are interpolated (lighter bars).

F Attribution rules

The product–difference identity (8) attributes $W_{t_1}R_{t_1} - W_{t_0}R_{t_0}$ as $W_{t_0}\Delta R + \Delta W R_{t_1}$, the base pivot; the inverted pivot $W_{t_1}\Delta R + \Delta W R_{t_0}$ is equally exact, as is the corresponding swap inside $\Delta(n_g v_g)$. The two differ only in which channel absorbs the cross terms $\Delta W \Delta R$, and for the three-factor product $n_g v_g \rho_g$ there are six orderings of the factors. The midpoint rule averages the two pivots at each level of the product; the Shapley rule averages over all orderings, which is the midpoint for a two-factor product but not for a three-factor one [Shorrocks, 2013]. Table 18 reports the four channels under the four rules for the main partitions. Under gender, the rules agree to within a few thousandths. Under education, where the basic-educated group's

population share and variance share both collapsed, the split of the re-weighting between the compositional and the inequality channel depends on the rule (−0.006 and −0.053 under the base pivot, −0.087 and +0.020 under Shapley), while their sum and the within channel also vary. The group-gap channel does not. For the fine partitions the sensitivity is larger. Under age×education the within channel is −0.017 under the base pivot, +0.028 under the inverted pivot and +0.004 under Shapley, and the compositional and inequality channels swap several hundredths between rules. The paper therefore reports the Shapley rule throughout, reasons with the sum of the two re-weighting channels, and reads the within channel of the fine partitions as small under every rule rather than as a precise number.

Table 18: Four channels under four attribution rules, 1986 → 2014

Partition	Attribution	$\Delta\rho$	Within	Compositional	Inequality	Group-gap
Gender						
	Base pivot	+0.065	+0.066	+0.022	+0.021	-0.044
	Inverted pivot	+0.065	+0.064	+0.026	+0.019	-0.044
	Midpoint	+0.065	+0.065	+0.024	+0.020	-0.044
	Shapley	+0.065	+0.065	+0.024	+0.020	-0.044
Age						
	Base pivot	+0.065	+0.019	+0.049	+0.065	-0.068
	Inverted pivot	+0.065	+0.031	+0.006	+0.096	-0.068
	Midpoint	+0.065	+0.025	+0.027	+0.081	-0.068
	Shapley	+0.065	+0.025	+0.027	+0.081	-0.068
Education						
	Base pivot	+0.067	+0.052	-0.006	-0.053	+0.074
	Inverted pivot	+0.067	+0.073	-0.162	+0.082	+0.074
	Midpoint	+0.067	+0.062	-0.088	+0.019	+0.074
	Shapley	+0.067	+0.060	-0.087	+0.020	+0.074
Age×*Education*						
	Base pivot	+0.067	-0.017	+0.055	+0.016	+0.013
	Inverted pivot	+0.067	+0.028	-0.139	+0.164	+0.013
	Midpoint	+0.067	+0.006	-0.045	+0.093	+0.013
	Shapley	+0.067	+0.004	-0.044	+0.094	+0.013
Gender×*Age*×*Education*						
	Base pivot	+0.067	-0.023	+0.066	+0.014	+0.010
	Inverted pivot	+0.067	+0.001	-0.071	+0.126	+0.010
	Midpoint	+0.067	-0.011	-0.002	+0.070	+0.010
	Shapley	+0.067	-0.011	-0.002	+0.070	+0.010

Notes: Decomposition of $\Delta\rho$ from 1986 to 2014 ($h = 5$) under the base pivot of equation (9), the inverted pivot, the hierarchical midpoint, and the Shapley rule. All four sum to $\Delta\rho$ exactly.

G Stylized facts by subgroup

This appendix documents the raw stylized facts that motivate the decomposition. For each year I plot three descriptive moments of income: the *level* (μ_t, the cross-sectional mean of the log real wage), the *inequality* (σ_t, its standard deviation), and the *persistence* (r_t, the correlation of the log wage across a five-year horizon), computed overall (Figure 20) and separately by

gender, age, industry and geography (Figures 21–24); for the subgroup partitions I additionally plot the *composition* (n_t, each group's share of employment), which drives the compositional channel of Section 3.2, and each group's *variance share* ($v_t = \text{Var}_g/\text{Var}$, its share of the total within-year rank variance), which together with n_t forms the subgroup weight $\omega_g = n_g v_g$ and drives the inequality channel. The heterogeneity visible here, in every dimension and in all three moments and in the shifting subgroup composition, is exactly what a single population mobility number conceals. Within-group persistence differs sharply across groups, dispersion shares diverge from population shares, and the mean gaps between groups that drive the gap term are large in some partitions (education, age) and small in others (region, industry). This is the raw material that Sections 3 and 4 organize into an exact accounting.

Appendix G.1 reports the measure-specific inputs to the four channels of the change decomposition discussed in Section 4.2. The *within-group persistence* panels (ρ_g for the rank-rank slope and r_g for the log-wage correlation) drive the *within* channel (blue). The within bar moves when the group persistence lines climb or fall. The *composition* panel (the "n_g, Population share of each group" panel) drives the *compositional* channel (green). Shifting employment toward a group moves the aggregate in proportion to that group's dispersion-weighted persistence. The *variance-share* panel (v_g) and the *dispersion-share* panel (d_g) show the inputs to the *inequality* channel (purple) for the rank-rank slope and log-wage correlation, respectively. When a group's dispersion grows relative to the pool, its share of the variance rises. The mean-gap, gap-ratio and gap-weight panels describe the *group-gap* channel (red). The mean gaps between groups, and how stably they persist, are what that bar measures. The following examples refer to the rank-rank decomposition. Under *gender*, the mean-gap panel shows the male–female gap narrowing (red mobility-enhancing), the variance-share panel that women's rank-variance share is higher, the persistence panel that women are stickier and that both sexes grew stickier (blue positive), and the composition panel the rise in the female employment share (green positive). Under *age*, aging tilts the composition and variance-share panels toward the persistent older group (green and purple positive) while the fixed cohorts'

mean gaps compress across each window (red mobility-enhancing). Under *education*, large and persistent between-education mean gaps gain weight (red positive), the shrinking basic group loses weight (green mobility-enhancing), every group grows stickier (blue positive). Reading a change figure is reading these panels at once.

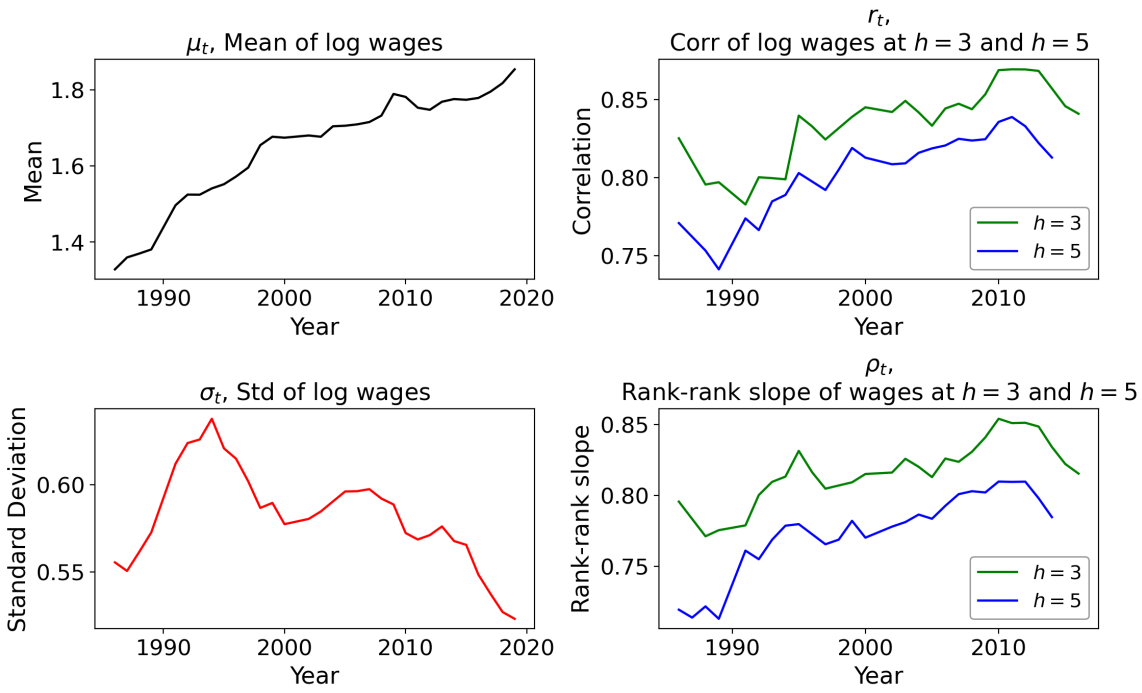


Figure 20: Stylized facts, Portugal 1986–2019: level, inequality, and persistence

Notes: μ_t and σ_t are the cross-sectional mean and population standard deviation of the log real wage each year (all wage earners). The persistence panels show the correlation of the log wage and the population rank-rank slope, each at lags $h = 3$ and $h = 5$. Years 1990 and 2001 are absent throughout. The survey was not conducted in those years.

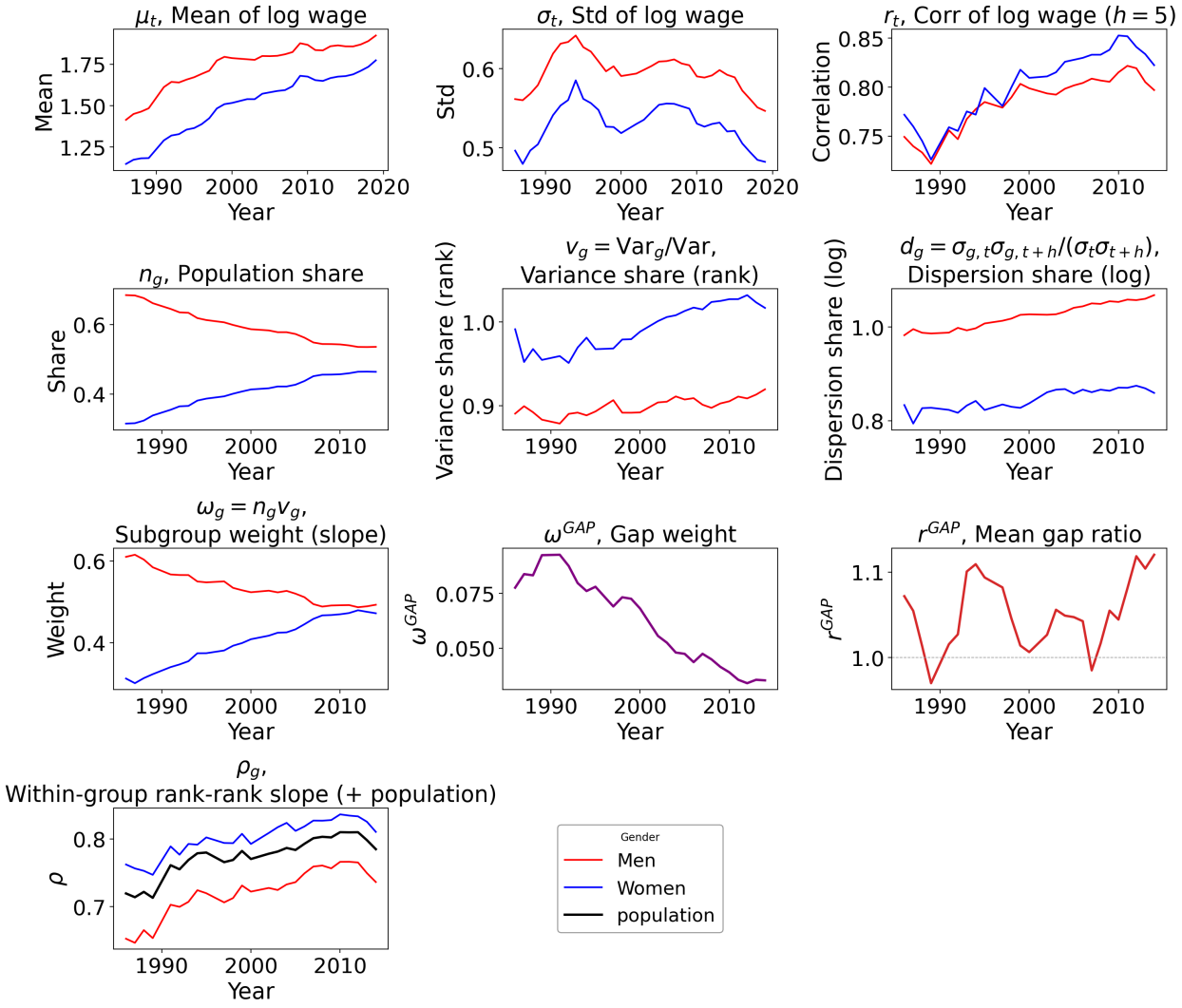


Figure 21: Level, inequality, persistence, and composition by gender

Notes: In Figures 21–24 each panel plots one series per group. *Row 1*: raw cross-sectional moments of the log wage: μ_t (mean), σ_t (standard deviation), and r_t (correlation across the five-year survivor pair). *Row 2*: composition and the two variance shares: n_g (population share); $v_g = \text{Var}_g/\text{Var}$, the share of the within-year *rank* variance (the rank-rank slope's inequality weight); and $d_g = \sigma_{g,t}\sigma_{g,t+h}/(\sigma_t\sigma_{t+h})$, the *log*-wage dispersion share (the correlation's dispersion weight). *Row 3*: the slope weights: subgroup weight $\omega_g = n_g v_g$, gap weight ω^{GAP}, and mean gap ratio r^{GAP}. *Row 4*: within-group rank-rank slopes ρ_g with the population slope ρ. Rows 2–4 are survivor, base-year quantities from the level decomposition of Section 3.1; row 1 is cross-sectional. A single group's variance or dispersion share can exceed one, since Var_g can exceed the pooled variance; only the population-weighted sum $\sum_g n_g v_g < 1$. Years 1990 and 2001 are absent (no survey).

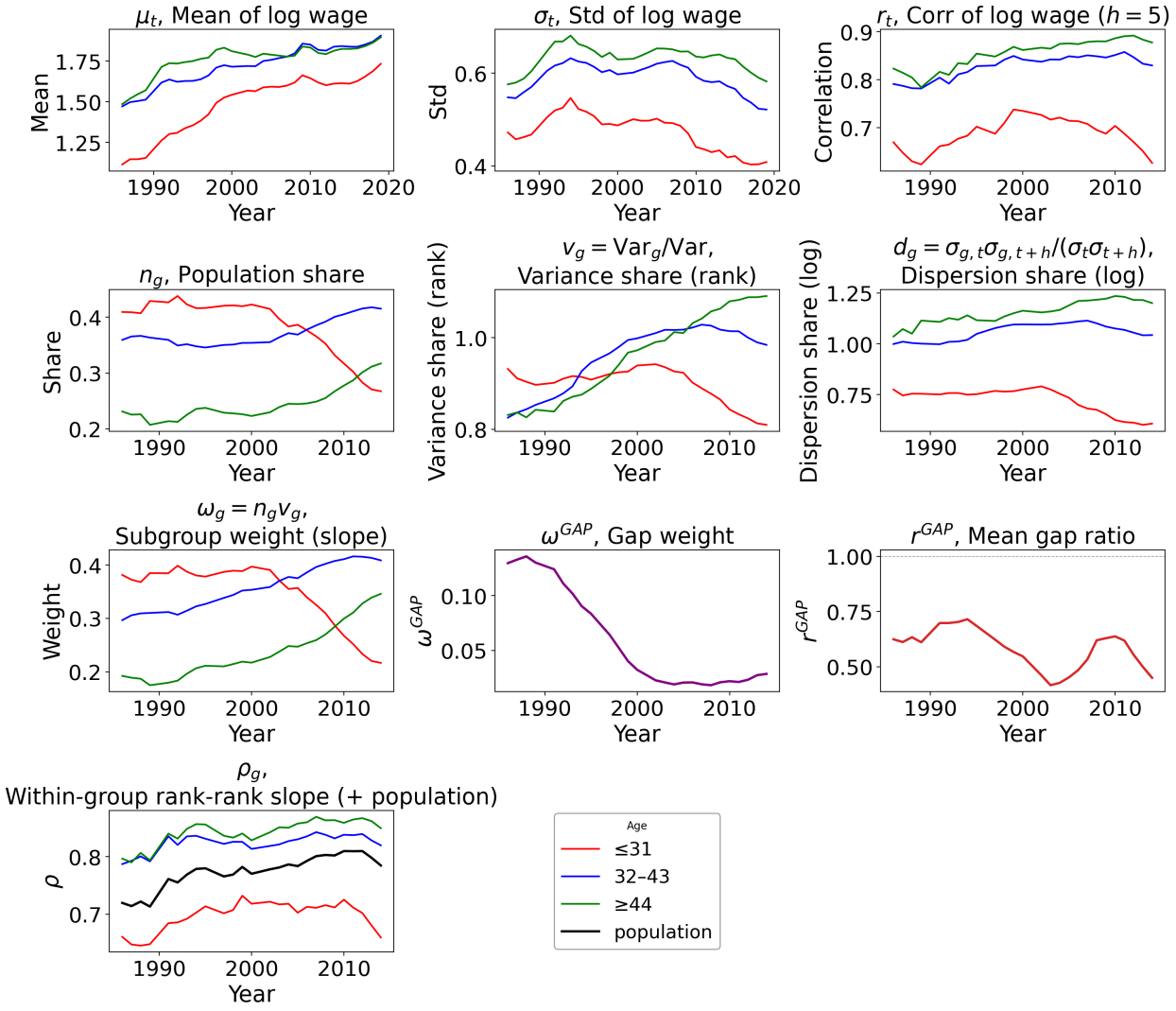


Figure 22: Level, inequality, persistence, and composition by age bracket

Notes: Years 1990 and 2001 are absent throughout. The survey was not conducted in those years.

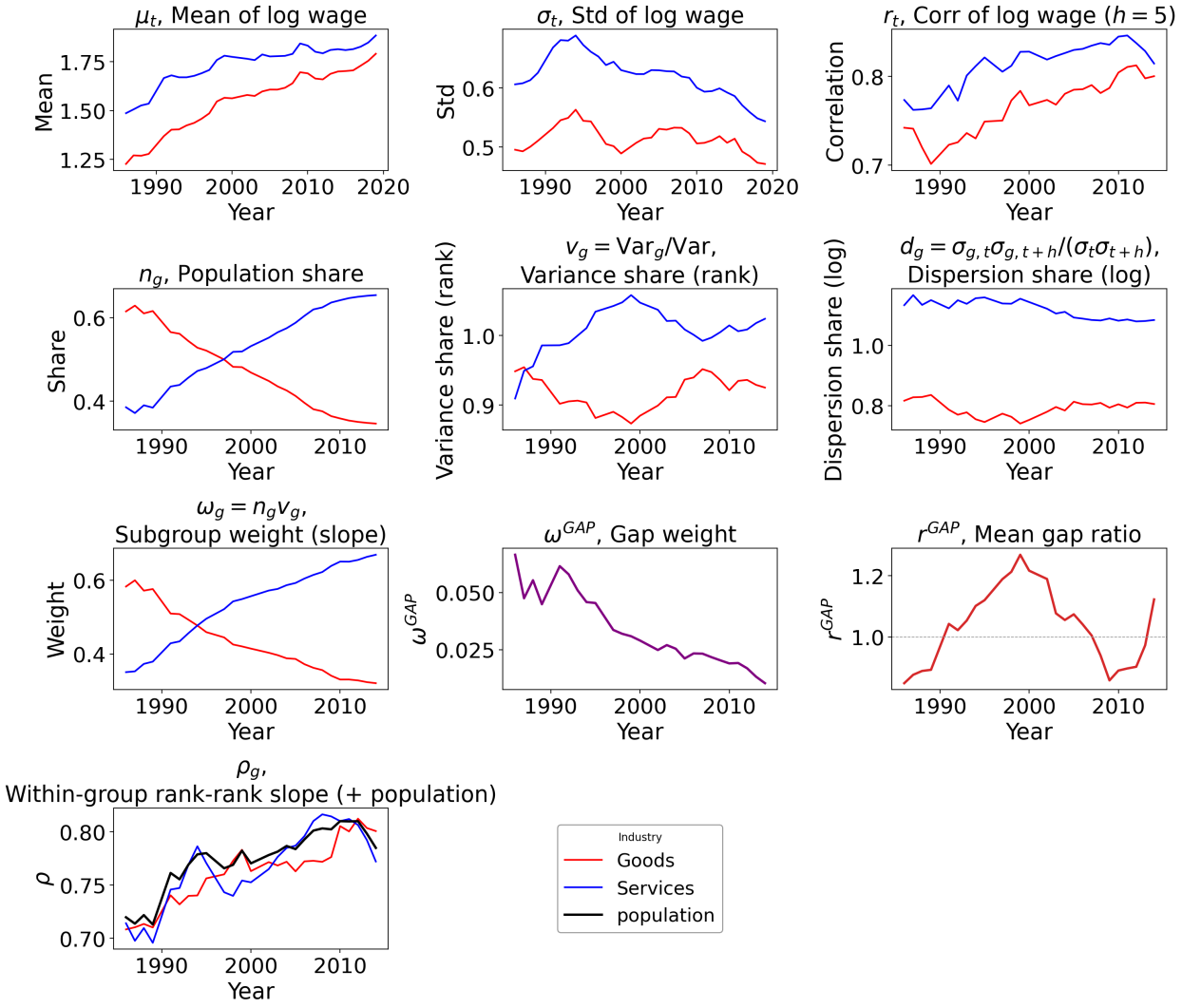


Figure 23: Level, inequality, persistence, and composition by broad industry (goods vs. services)

Notes: Years 1990 and 2001 are absent throughout. The survey was not conducted in those years.

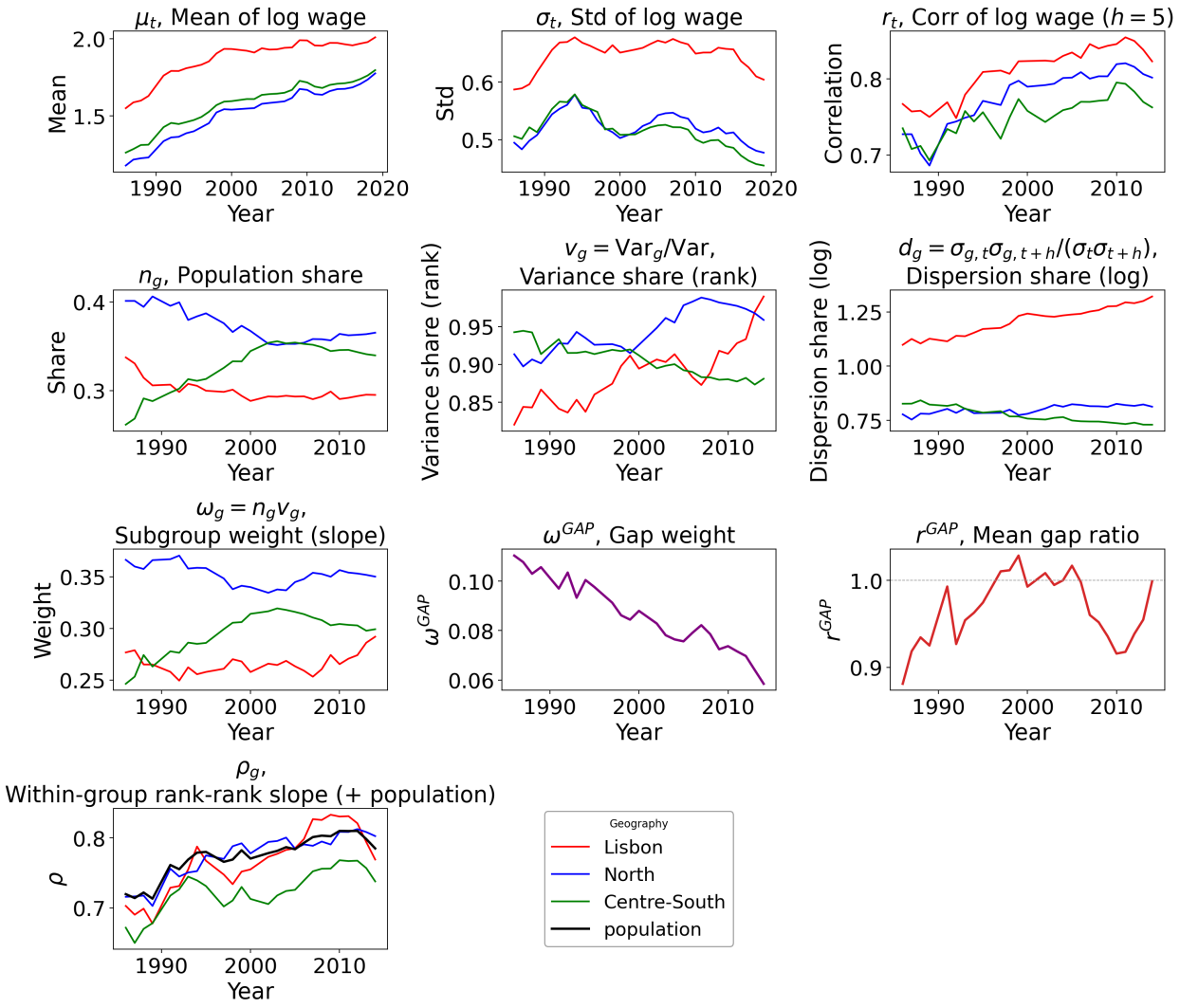


Figure 24: Level, inequality, persistence, and composition by broad geography (Lisbon / North / Centre–South)

Notes: Years 1990 and 2001 are absent throughout. The survey was not conducted in those years.

G.1 Decomposition terms over time, by measure

The figures above summarize the raw moments and the slope-decomposition ingredients together. This subsection instead plots *every* term of the level decomposition over base years, separately for each measure, the rank-rank slope (Figures 25–29) and the log-wage correlation (Figures 30–34), and for each partition. Each panel grid shows the within-group persistence ρ_g (or r_g) with the population value; the subgroup weights ω_g with their total $\sum_g \omega_g$; the composition shares n_g; the variance share v_g (rank) or dispersion share d_g (log); the

mean gap ratio r^{GAP}; the base-year mean gaps between groups $\bar{y}_t^{GAP}$ by pair; and the factorisation of the gap weight into composition balance $n_g n_k$ and squared standardized gap $(\bar{y}_t^{GAP}/\sigma)^2$, whose pairwise product summed is the gap weight ω^{GAP}. Together they trace how $\rho = \sum_g \omega_g \rho_g + \omega^{GAP} r^{GAP}$, and its correlation analogue, is assembled and how each piece evolves.

G.1.1 Rank-rank slope

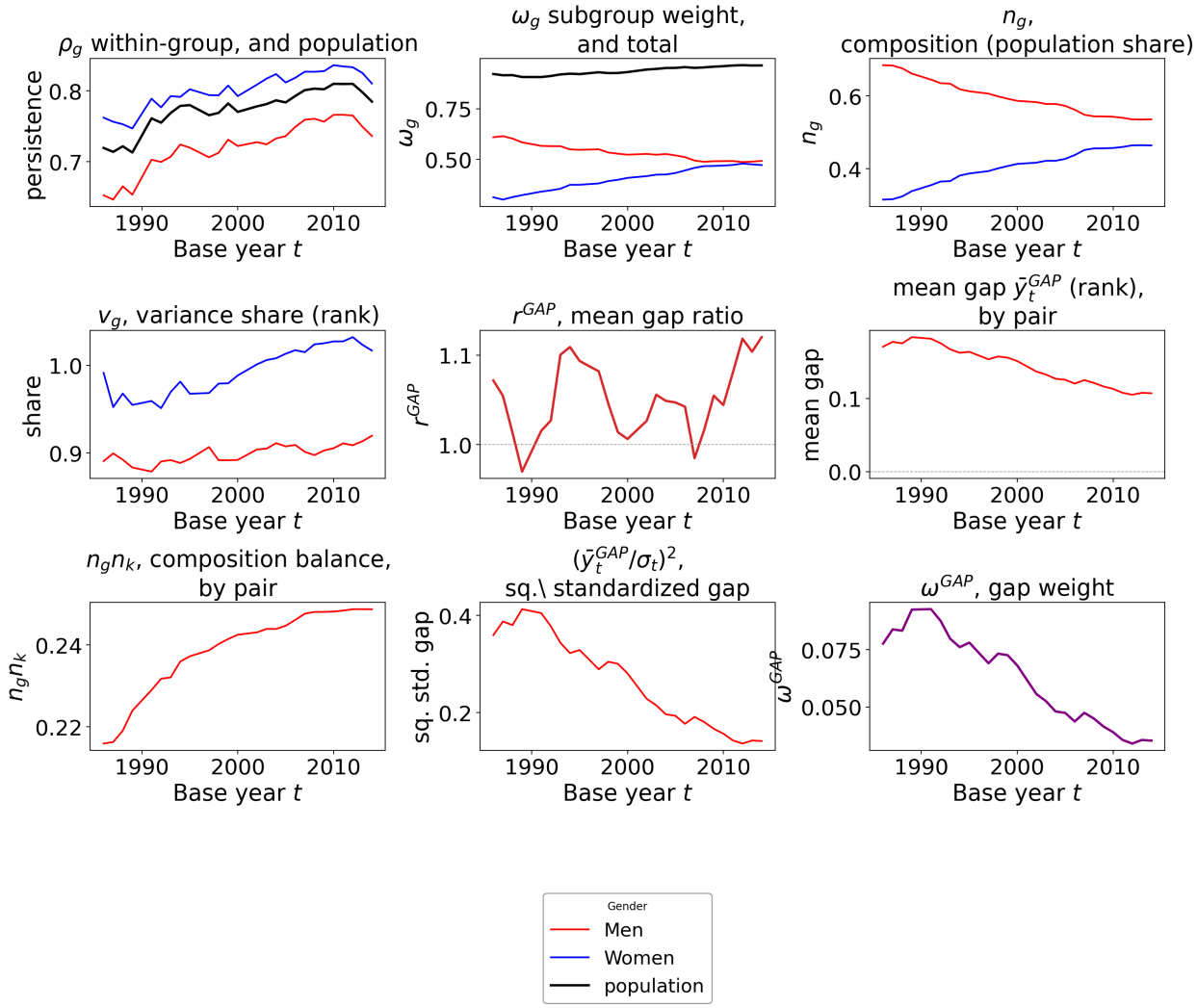


Figure 25: Slope decomposition terms over time, gender

Notes: Level-decomposition terms of the rank-rank slope (Section 3.1) at horizon $h = 5$, by base year t. Panels, by row: within-group slopes ρ_g with the population slope ρ; subgroup weights $\omega_g = n_g v_g$ with their total $\sum_g \omega_g$; population shares n_g; rank-variance shares $v_g = \text{Var}_g/\text{Var}$; mean gap ratio r^{GAP}; base-year mean rank gaps $\bar{y}_t^{GAP}$ by pair; and the gap-weight factorisation: composition balance $n_g n_k$, squared standardized gap $(\bar{y}_t^{GAP}/\sigma_t)^2$, and the gap weight $\omega^{GAP} = \sum_{g<k} n_g n_k (\bar{y}_t^{GAP}/\sigma_t)^2$ (their pairwise product, summed). Base years without a five-year pair are interpolated.

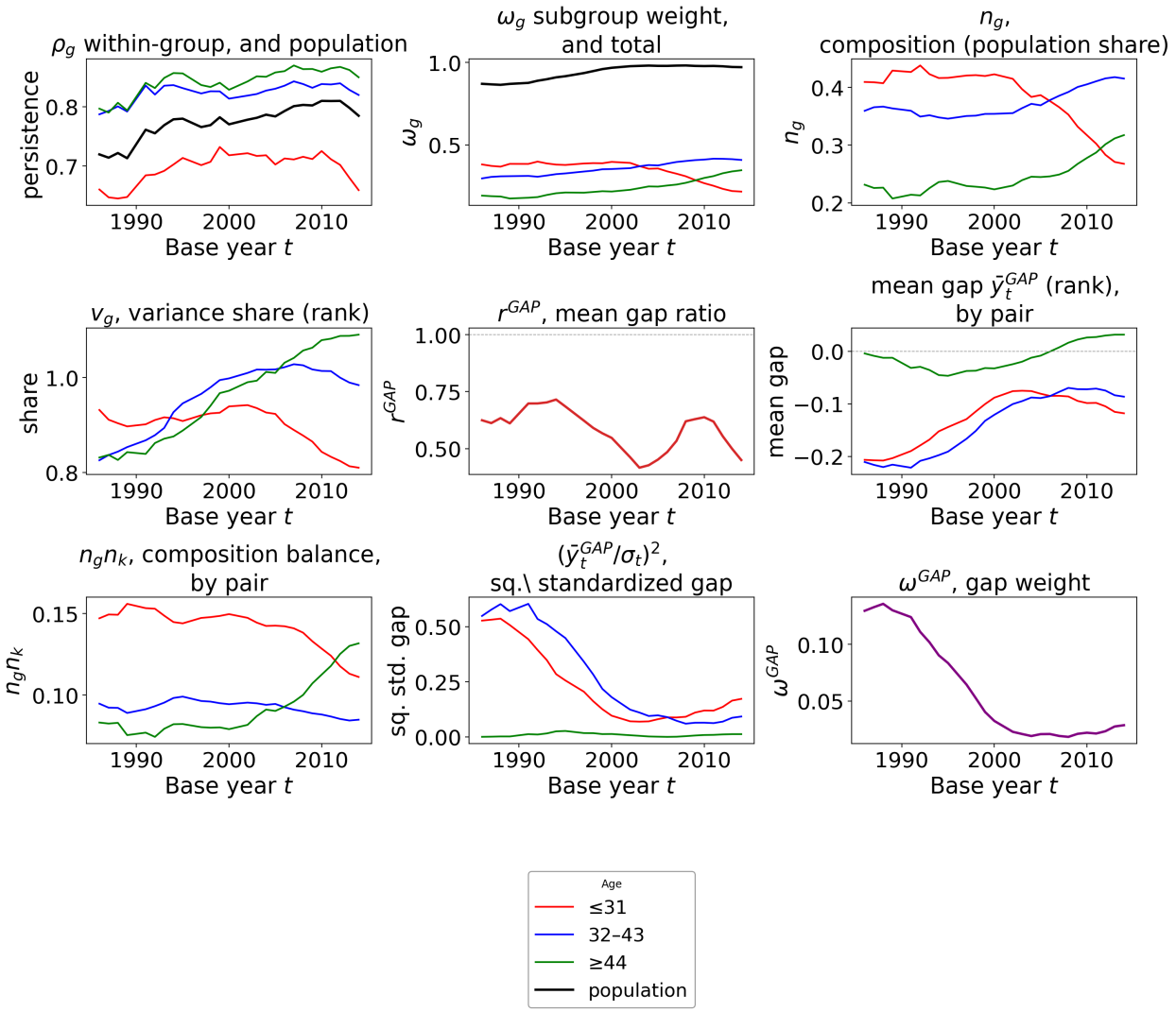


Figure 26: Slope decomposition terms over time, age

Notes: As in Figure 25, for the age partition (equal-share terciles).

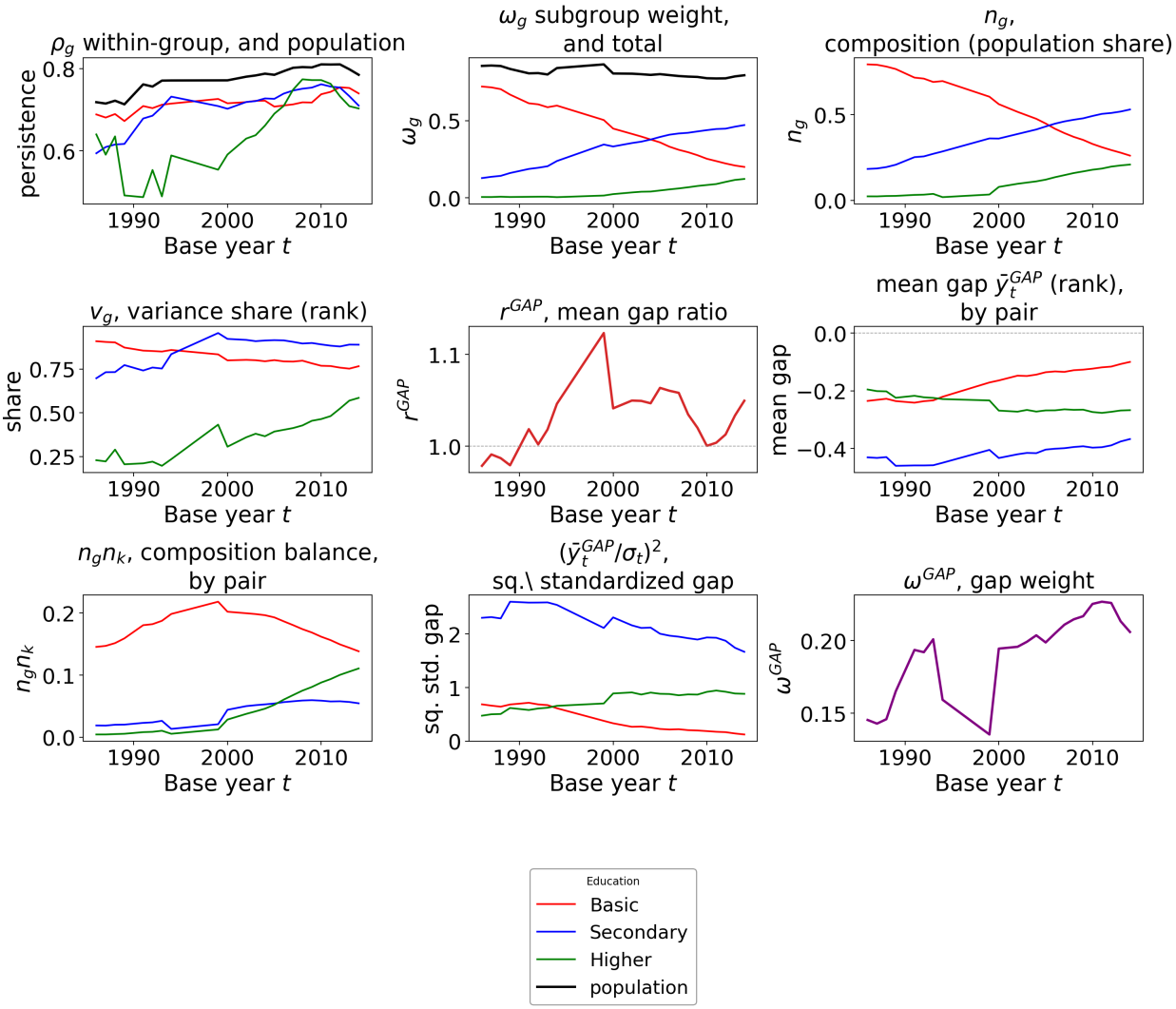


Figure 27: Slope decomposition terms over time, education

Notes: As in Figure 25, for the education partition. Education is unrecorded in 1995, 1997, and 1998.

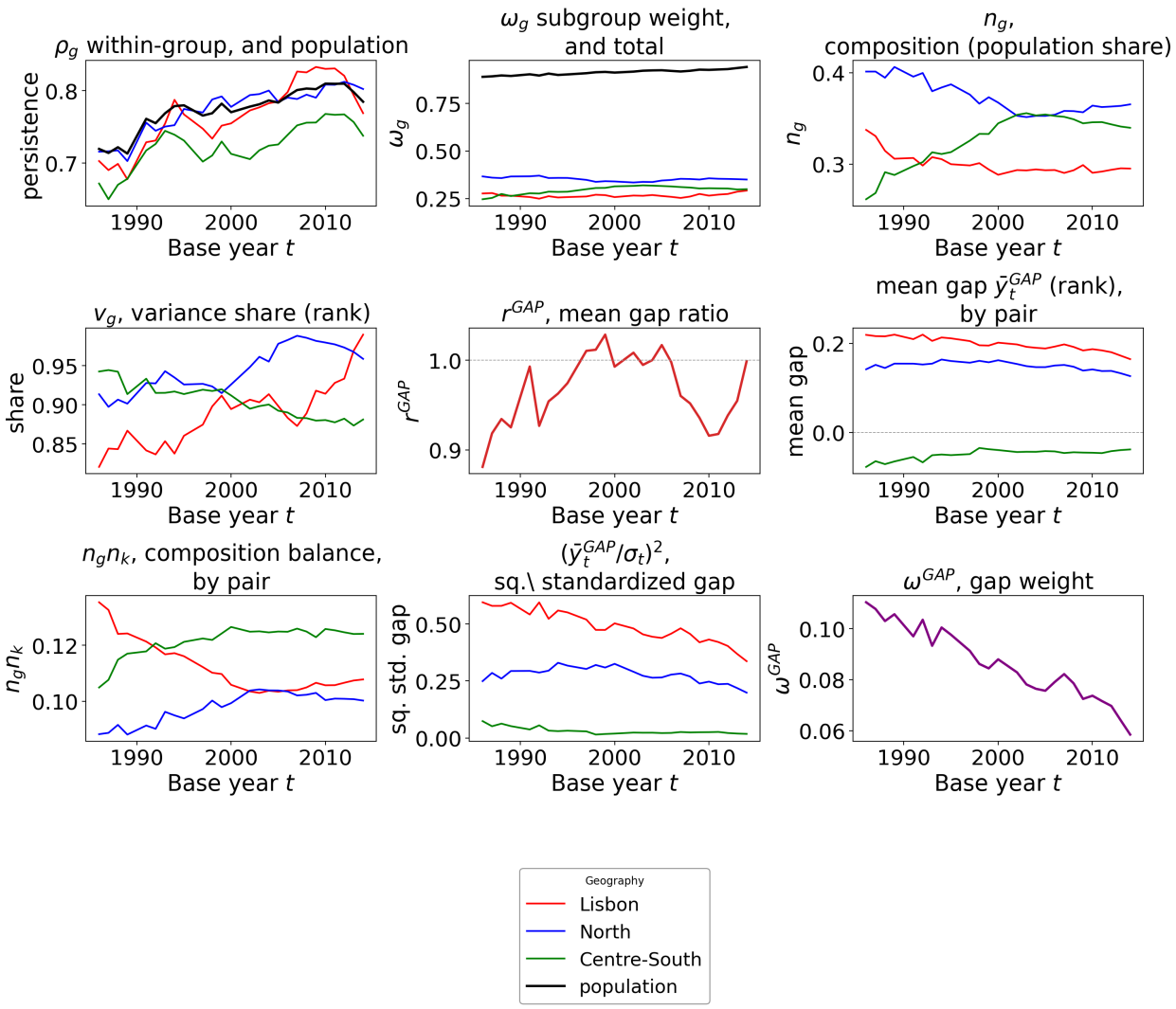


Figure 28: Slope decomposition terms over time, geography

Notes: As in Figure 25, for the broad-geography partition (Lisbon / North / Centre–South).

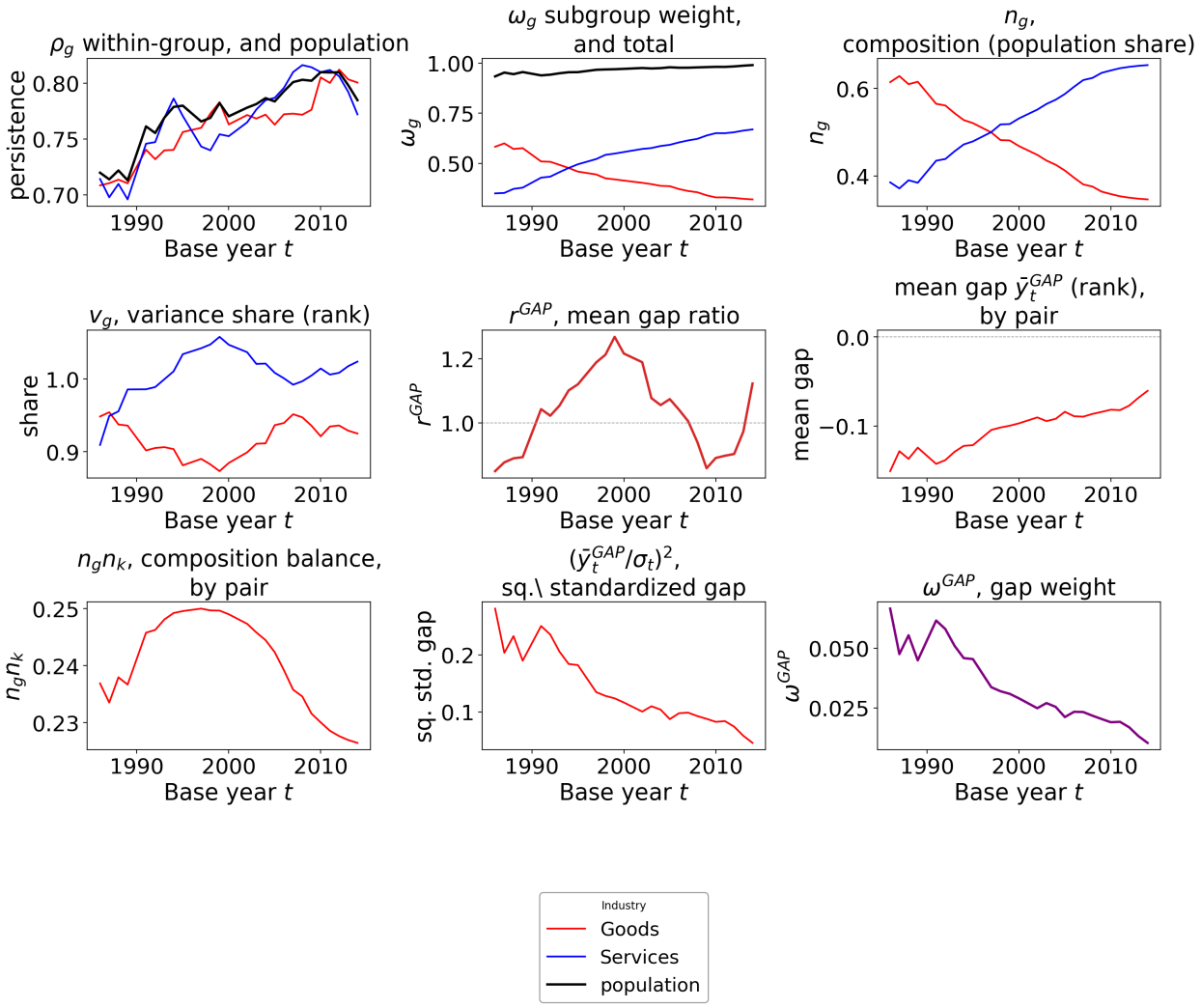


Figure 29: Slope decomposition terms over time, broad industry

Notes: As in Figure 25, for the broad-industry partition (Goods / Services).

G.1.2 Correlation of log wages

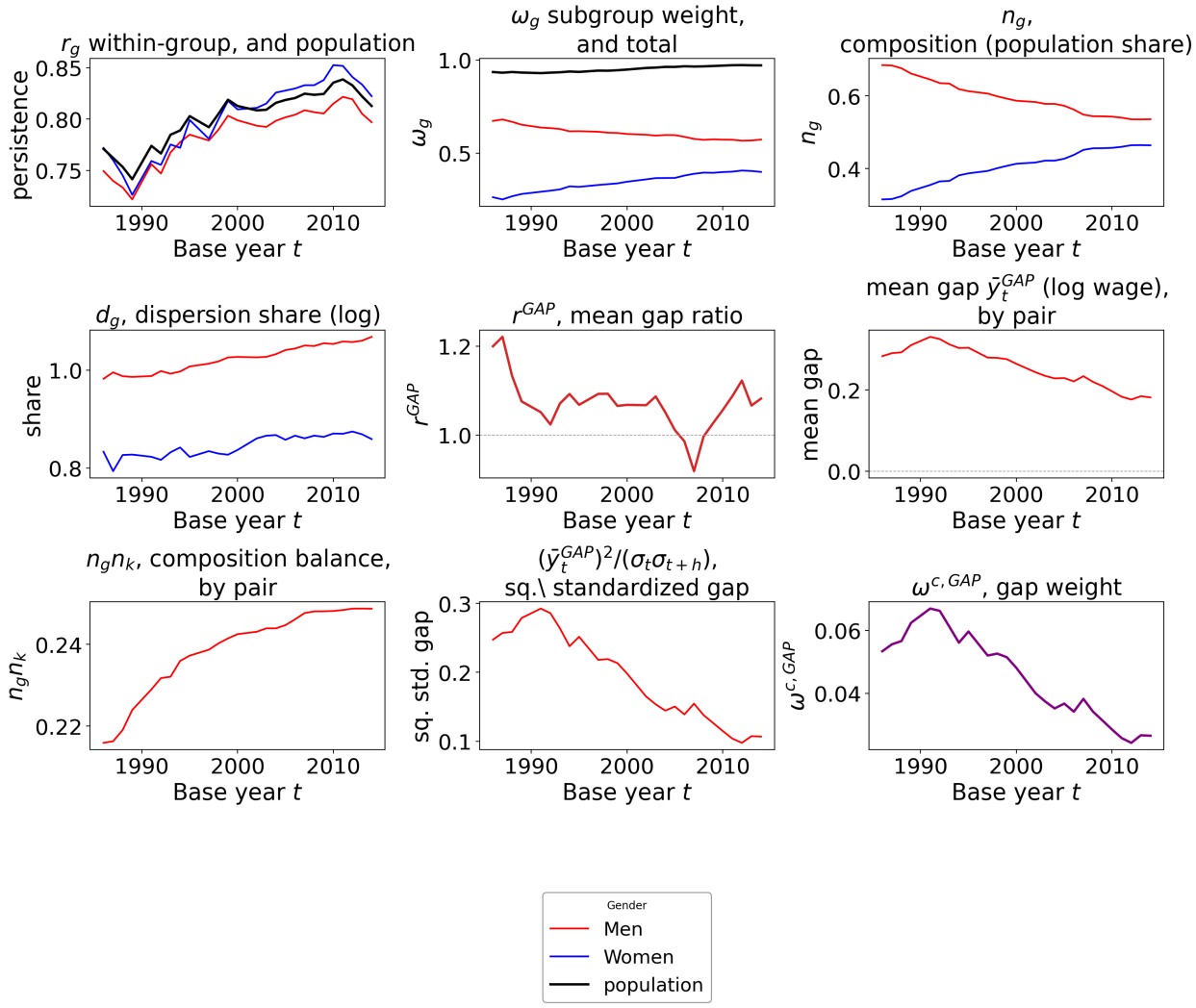


Figure 30: Correlation decomposition terms over time, gender

Notes: Level-decomposition terms of the log-wage correlation (Appendix B) at horizon $h = 5$, by base year t. Panels, by row: within-group correlations r_g with the population correlation r; subgroup weights $\omega_g^c = n_g d_g$ with their total; population shares n_g; dispersion shares $d_g = \sigma_{g,t}\sigma_{g,t+h}/(\sigma_t\sigma_{t+h})$; mean gap ratio r^{GAP}; base-year mean log-wage gaps $\bar{y}_t^{GAP}$ by pair; and the gap-weight factorisation: composition balance $n_g n_k$, squared standardized gap $(\bar{y}_t^{GAP})^2/(\sigma_t\sigma_{t+h})$, and the gap weight $\omega^{c,GAP} = \sum_{g<k} n_g n_k (\bar{y}_t^{GAP})^2/(\sigma_t\sigma_{t+h})$. Base years without a five-year pair are interpolated.

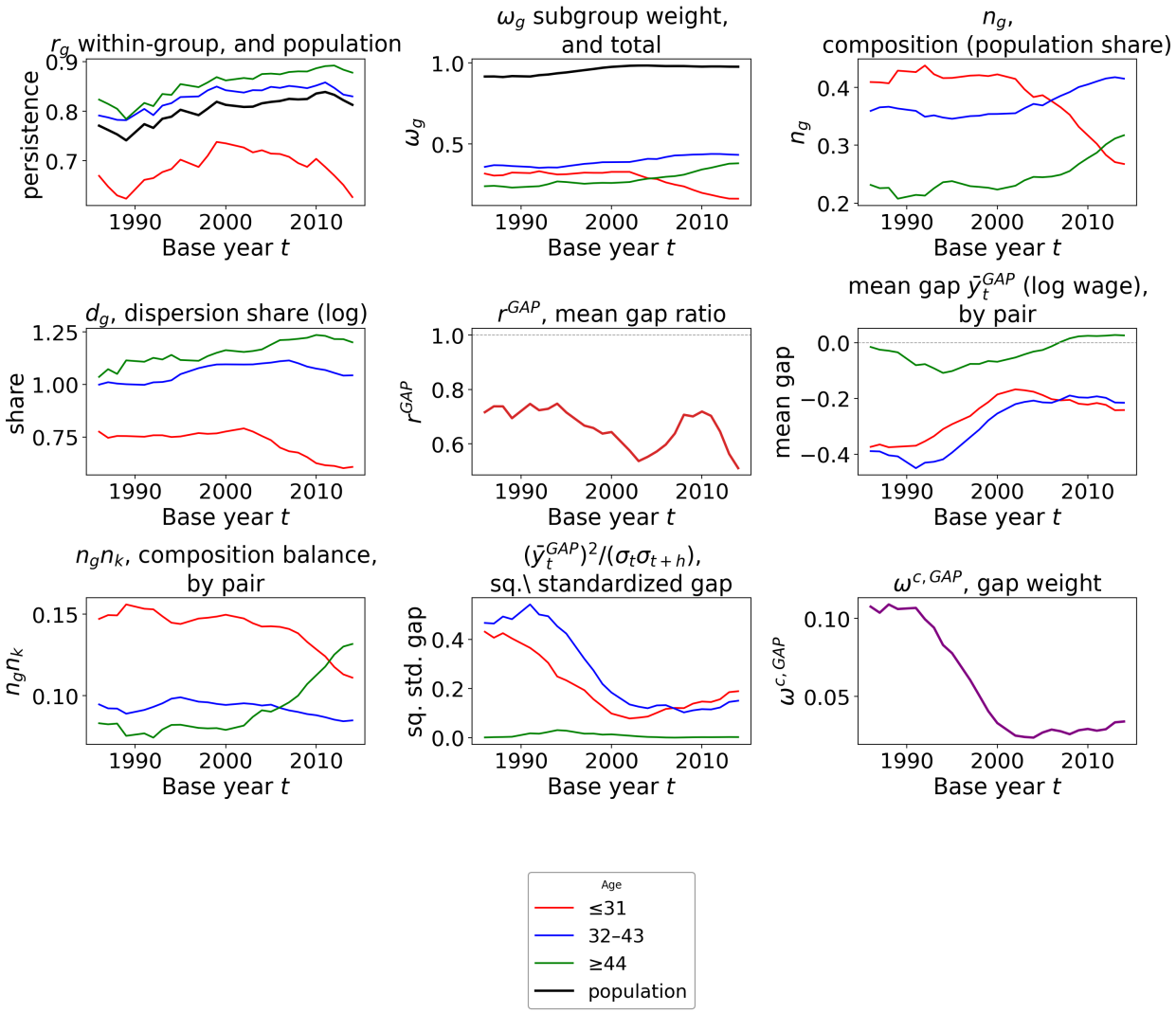


Figure 31: Correlation decomposition terms over time, age

Notes: As in Figure 30, for the age partition (equal-share terciles).

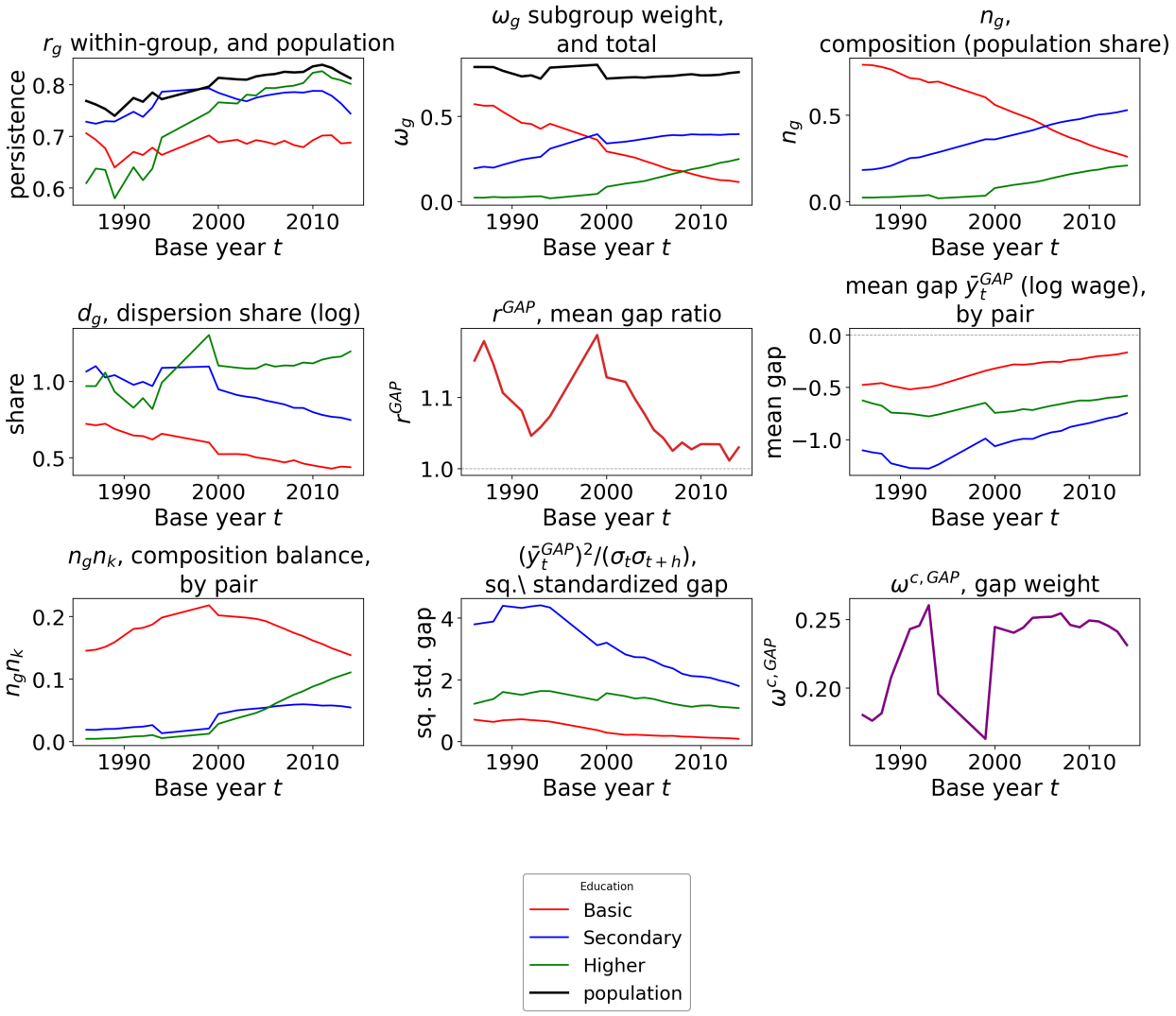


Figure 32: Correlation decomposition terms over time, education

Notes: As in Figure 30, for the education partition. Education is unrecorded in 1995, 1997, and 1998.

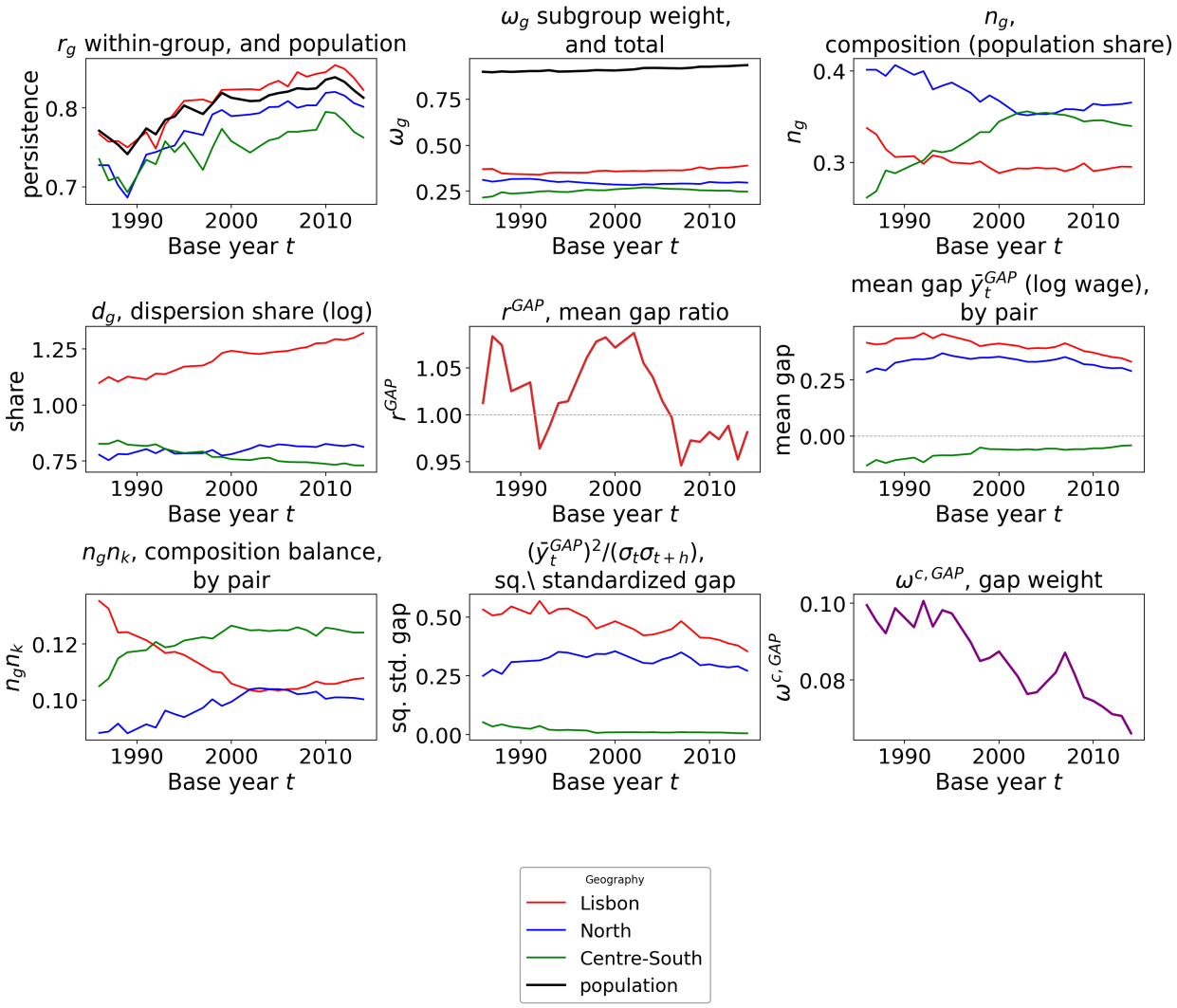


Figure 33: Correlation decomposition terms over time, geography

Notes: As in Figure 30, for the broad-geography partition (Lisbon / North / Centre–South).

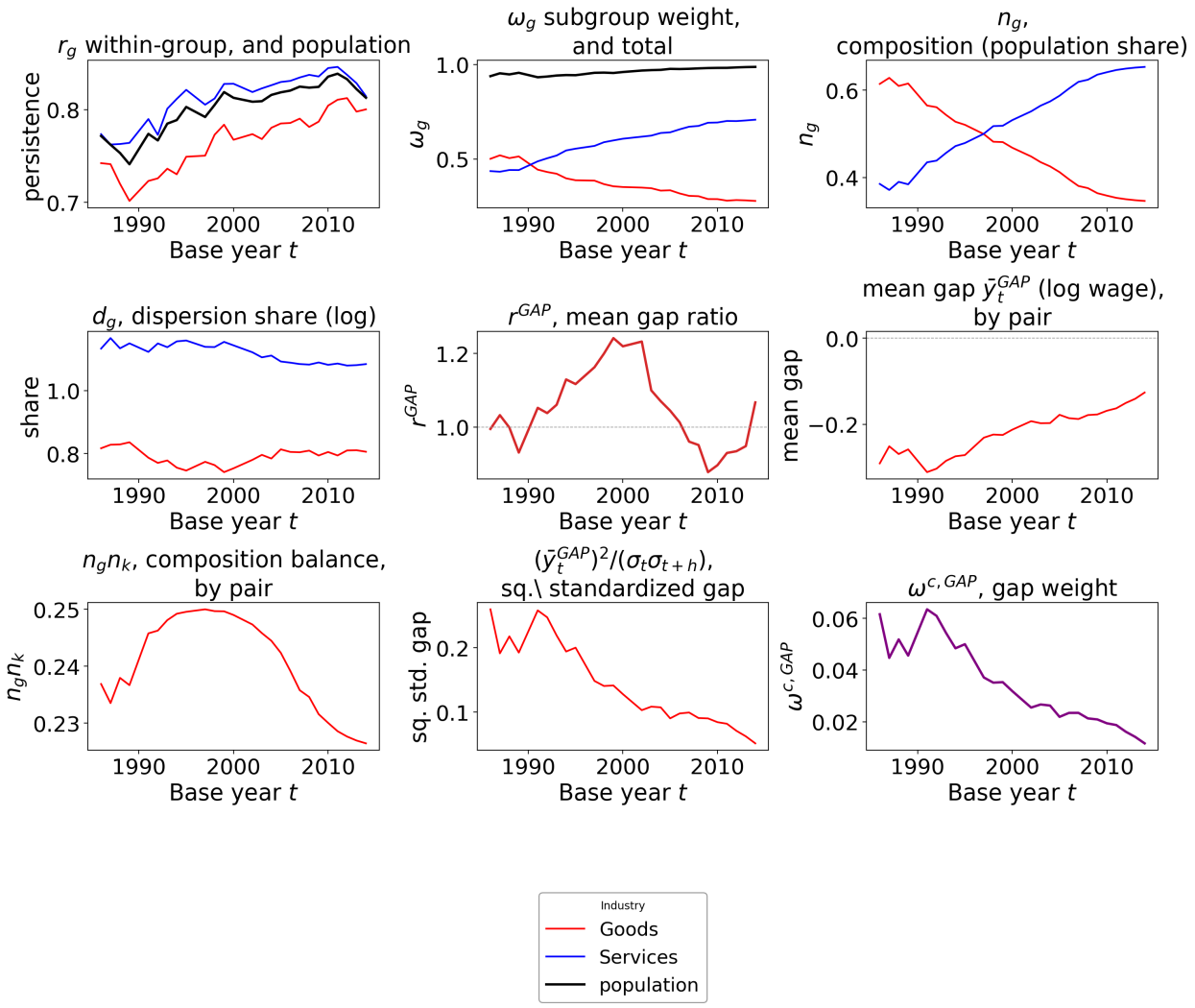


Figure 34: Correlation decomposition terms over time, broad industry

Notes: As in Figure 30, for the broad-industry partition (Goods / Services).

H Cross-country context and Great Gatsby curves

The main text documents a pronounced decline in Portuguese intragenerational mobility. Figure 35 places it in cross-country context using the harmonized Global Repository of Income Dynamics [Guvenen et al., 2022]. Across the thirteen GRID countries, intragenerational mobility, measured by *absolute upward mobility* (AUM), the expected five-year-ahead rank of a below-median worker, is essentially flat over 1987–2007, with large differences in level (the Nordic countries most mobile; the UK, Italy, and Brazil least). Computed on the same defini-

tion from the Portuguese data, for workers aged 25–55 as in GRID, the Portuguese series falls by around four percentile points over the same window, from 34.1 in 1987 and 34.2 in 1997, the level of Sweden (34.2 and 35.0) and above Denmark and Canada, to 29.8 in 2007, below Germany (30.1) and Mexico (30.0) and just above France (29.7), a movement that exceeds almost every cross-country gap in the figure. Whatever drives the Portuguese decline, it is not a force common to these economies. The Portuguese series ranks the three-year average of the log hourly wage over the years available among $t-1$, t and $t+1$, within the 25–55 population of each year; computed on all ages it runs two to three points higher, because young workers starting below the median rise fastest, and falls by six points. GRID's other headline statistic, the rank-rank slope of permanent-income ranks pooled over 1997–2007, is 0.86 for Portugal on the same window, the level of Italy (0.86) and just below the United Kingdom and Brazil (0.87), the least mobile economies in GRID; Portugal's pooled absolute upward mobility over the same years, 32.0, is mid-table (Canada 32.2, United States 31.7), the two statistics differing because the slope summarizes the whole distribution and upward mobility only its lower half. GRID has since added Portugal itself, on *Quadros de Pessoal* from 2002, which will allow a like-for-like comparison.

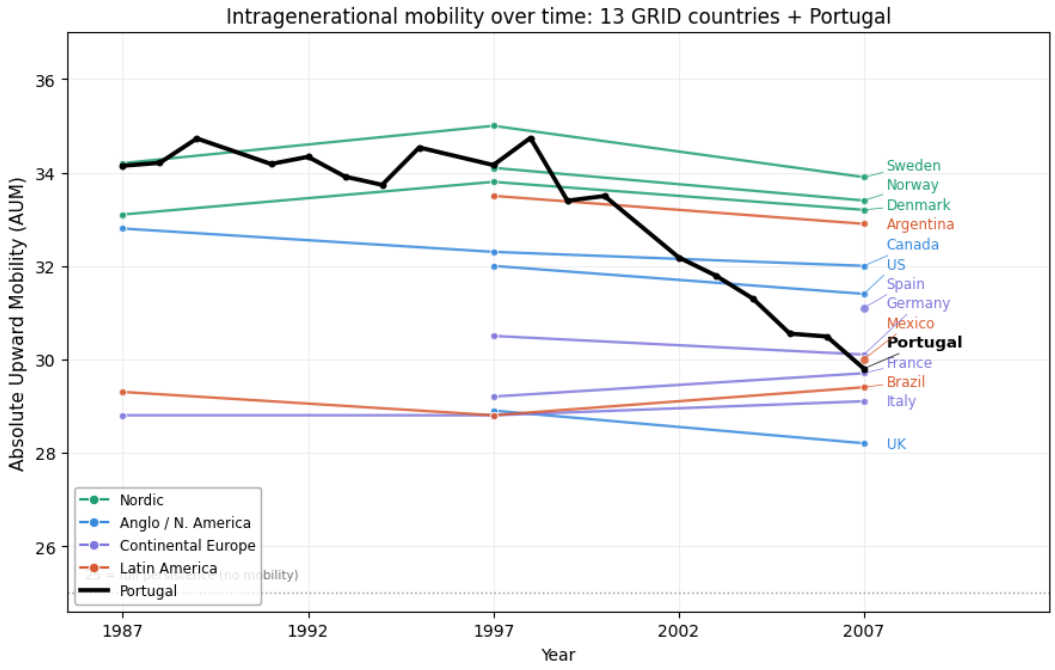


Figure 35: Intragenerational mobility over time: GRID countries and Portugal

Notes: Absolute upward mobility, AUM $= \mathbb{E}[\text{rank}_{t+5} \mid \text{rank}_t \leq 50]$, the expected five-year-ahead national percentile rank of a worker starting below the median; higher values indicate more mobility (25 = full persistence, 50 = no persistence). GRID values for 1987/1997/2007 are from Guvenen et al. [2022], Table 4; the Portuguese series is computed identically from the Quadros de Pessoal on permanent-income ranks, for workers aged 25–55 at the base year, ranked within that age group. Because GRID uses annual earnings rather than hourly wages, the two sources are not strictly comparable in *level*, but the time pattern is.

The cross-country "Great Gatsby curve" records a negative association between inequality and mobility. More unequal economies are less mobile [Guvenen et al., 2022]. The level decomposition suggests an intragenerational, within-economy analogue. Because a subgroup's imprint on population mobility runs through the weight $\omega_g = n_g \sigma_g^2/\sigma^2$, a group's *own* inequality is already a first-order determinant of how it moves the aggregate; the question is whether more unequal subgroups are also *internally* less mobile.

For each of the paper's partitions I plot, for every subgroup and base year ($h = 5$, 1986–2014), base-year inequality, the standard deviation of the log real wage within the subgroup-year, against immobility. Workers are ranked *within their own subgroup* each year, so cross-subgroup inequality enters only through the horizontal axis. Figure 36 shows a clear positive relationship. More unequal subgroups are markedly less mobile (Pearson $r = 0.51$ on the

rank-rank slope across all subgroup-years, and 0.62 once subgroup-years are weighted by their population, so that small groups do not pull the fit). The gradient is of similar strength for the rank-rank slope and for the log-income elasticity (0.46 unweighted, 0.55 weighted), and the comparison between the two is informative. The elasticity equals the correlation times the dispersion ratio σ_{t+h}/σ_t, so it mechanically carries the very inequality plotted on the horizontal axis; the rank-rank slope normalizes the marginals away. Although inequality is not the object of this paper, the gradient *within* each partition is also informative. It is steep for age (0.89 on the rank-rank slope), present for geography (0.69) and industry (0.43), and flat for education on the rank-rank slope (−0.13 unweighted, 0.12 weighted) though positive on the elasticity (0.52). Gender is the exception. More dispersed men are, if anything, slightly *more* mobile than less dispersed women, so its gradient is flat for the rank-rank slope and clearly negative for the elasticity (−0.53), mirroring the compressing gender gap that the level decomposition flags. The usual Great Gatsby logic expects inequality and immobility to move together, but that holds only when the extra inequality reflects permanent differences between workers. If instead the wider dispersion comes from transitory shocks, larger year-to-year fluctuations rather than more durable gaps, the two need not be associated in that way, and can even move apart. Appendix H.1 makes this precise. Read through that lens, the gender reversal is suggestive. Men's greater inequality may be more transitory in nature, so their wider distribution need not translate into greater persistence. I put this forward only as a possibility, as the underlying income process is not something I identify here.

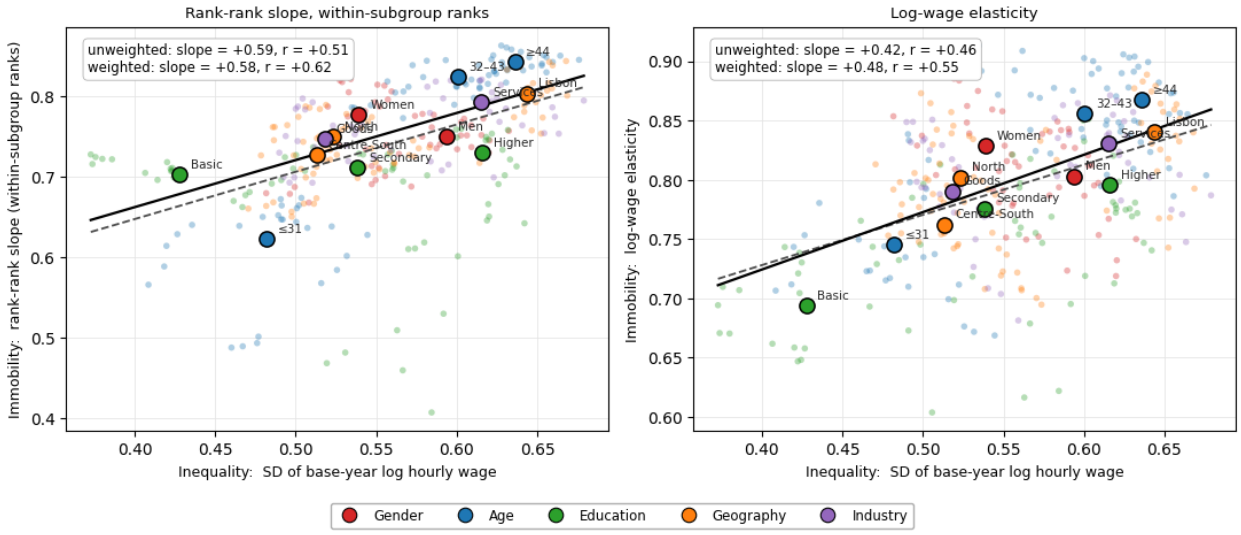


Figure 36: A Great Gatsby curve across subgroups (within-subgroup ranks, $h = 5$)

Notes: Each dot is a subgroup in one base year, over the five partitions of Section 2.3; larger ringed dots are population-weighted centroids. The horizontal axis is within-subgroup-year inequality (population standard deviation of the log real wage). *Left:* immobility measured by the rank-rank slope, with workers ranked within their own subgroup each year. *Right:* the log-income elasticity. Dashed lines are unweighted OLS fits across all subgroup-years and solid lines population-weighted fits; the slopes and Pearson r are annotated.

The intertemporal version of the curve does not hold. Figure 37 plots, across base years, the cross-sectional dispersion of log wages against population persistence. The two are essentially unrelated. The correlation is 0.23 for the rank-rank slope and −0.03 for the log-wage correlation (unweighted), and 0.06 and −0.18 when base years are weighted by sample size. Wage dispersion in Portugal rose into the mid-1990s and fell back, while persistence rose until 2010–2012; whatever raised inequality in the 1990s did not freeze positions, which is what Appendix H.1 would predict if that inequality was largely transitory.

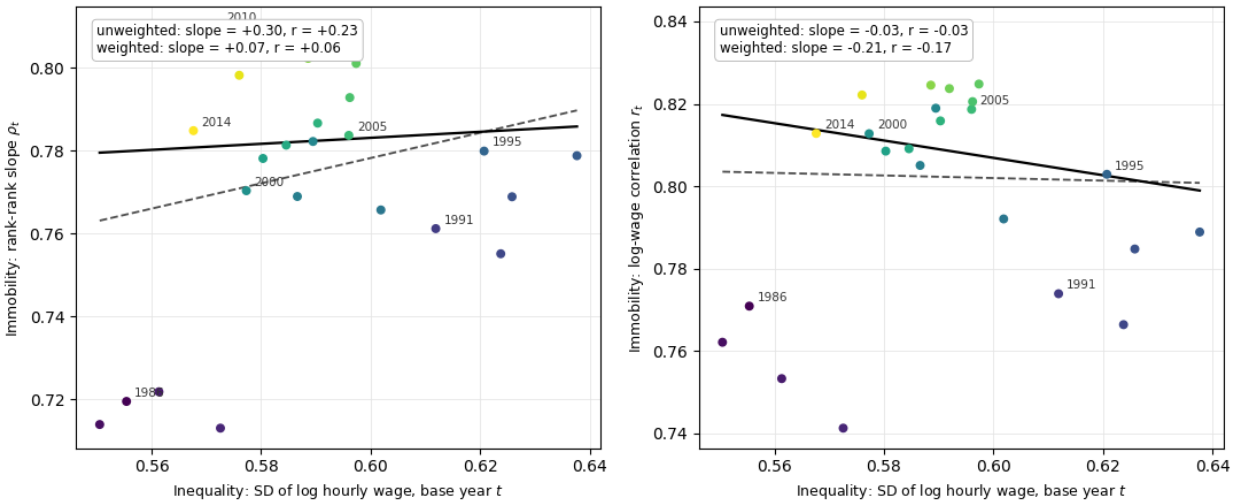


Figure 37: An intertemporal Great Gatsby curve for Portugal, 1986–2014

Notes: Each dot is a base year. The horizontal axis is the population standard deviation of the log real wage in the base year (all wage earners); the vertical axis is the five-year rank-rank slope (left) or the correlation of log wages (right). Dashed lines are unweighted fits, solid lines are fits weighted by the number of matched workers.

Two caveats. First, ranking *within* the subgroup rather than on the national ladder is deliberate. National ranks compress a narrow, high subgroup into a thin band of the distribution, so restriction of range mechanically deflates its measured slope; peer ranks remove this artifact and let each group's genuine internal persistence show. Second, the fit pools subgroup-years that are not independent, so it is descriptive, a subgroup-clustered standard error is the right object for inference. Set against the cross-country evidence of Figure 35, where intragenerational mobility is flat over time while Portugal's declines, the subgroup Gatsby curve reinforces the paper's reading. Inequality and immobility move together not only across economies but across the groups that compose one, though not, within Portugal, across periods.

H.1 A permanent–transitory income process

The Great Gatsby logic expects inequality and immobility to move together. Whether they do depends on where the extra inequality comes from. Consider the simplest permanent–transitory process for log income,

$$y_{i,t} = \alpha_i + \varepsilon_{i,t}, \tag{21}$$

with a permanent component α_i of variance σ^2_α and a transitory component $\varepsilon_{i,t}$, serially uncorrelated, independent of α_i, of variance σ^2_ε. Cross-sectional inequality is $\text{Var}(y_{i,t}) = \sigma^2_\alpha + \sigma^2_\varepsilon$, and the covariance between incomes h years apart is $\text{Cov}(y_{i,t+h}, y_{i,t}) = \sigma^2_\alpha$. The log-income slope (the intragenerational elasticity) and, since the process is stationary, the correlation coefficient both equal the share of inequality that is permanent:

$$\rho_{t,h} = r_{t,h} = \frac{\sigma^2_\alpha}{\sigma^2_\alpha + \sigma^2_\varepsilon}. \tag{22}$$

Inequality and immobility therefore co-move only when the additional dispersion is permanent. A rise in σ^2_α raises both the variance and the persistence of income, whereas a rise in σ^2_ε raises the variance and lowers persistence. A population, a period, or a subgroup can thus be more unequal and more mobile at once, if its extra dispersion is transitory. This is the lens through which this appendix reads the two exceptions to the Great Gatsby pattern in the Portuguese data: the flat relationship across periods, and the gender reversal, with more dispersed men being if anything more mobile than less dispersed women. The process is illustrative, not a claim about the true income process; the rank-rank slope used in the main text does not map into (22) one for one, since ranks remove the marginal distributions, but the same logic carries over. Transitory dispersion reshuffles ranks, permanent dispersion freezes them.