# Intergenerational Income Mobility in France: A Comparative and Geographic Analysis*

**Gustave Kenedi**
*Sciences Po*

**Louis Sirugue**
*Paris School of Economics*

This Version: August 2022

**Abstract**

We provide new estimates of intergenerational income mobility in France for children born in the 1970s using rich administrative data. Since parents' incomes are not observed, we employ a two-sample two-stage least squares estimation. Our results show that France is characterized by a strong persistence relative to other developed countries. 10% of children born to parents in the bottom 20% reach the top 20% in adulthood, four times less than children from the top 20%. We uncover substantial spatial variations in intergenerational mobility across departments, comparable to those observed across countries. We find that the upward mobility gains from geographic mobility are slightly decreasing in parent income and increasing in the income level of the destination department. The expected income rank of individuals from the bottom of the parent income distribution who moved towards high-income departments is around the same as the expected income rank of individuals from the 75$^{\text{th}}$ percentile who stayed in their childhood department.

**JEL Codes:** J62, R23, J61, C18

**Keywords:** intergenerational mobility, geographic mobility, spatial variations

*Kenedi: Sciences Po, 28 rue des Saints-Pères, 75007 Paris France, gustavekenedi@gmail.com. Sirugue: Paris School of Economics, 48 boulevard Jourdan, 75014 Paris France, louis.sirugue@psemail.eu, EHESS, CNRS, CI-Migration, INED. This paper was previously circulated under the title "The Anatomy of Intergenerational Income Mobility in France and its Spatial Variations". This work has been funded by public grants overseen by the French National Research Agency as part of the "Investissements d'avenir" program (references ANR-17-EURE-001 and ANR-10-EQPX-17 - Centre d'accès sécurisé aux données - CASD)), the ANR JOCE (reference ANR-18-CE41-0003-02), and the French Collaborative Institute on Migration coordinated by the CNRS (reference ANR-17-CONV-0001). This paper benefited from helpful feedback and suggestions from Pierre Cahuc, Pascale Champalaune, Pierre-Philippe Combes, Clément Dherbécourt, Basile Grassi, Thomas Piketty, Benoît Schmutz, Michaël Sicsic, Patrick Simon, Yann Thommen, and Alain Trannoy, as well as participants at the 2021 Augustin Cournot Doctoral Days, the 69$^{\text{th}}$ Congress of the French Economic Association, and seminars at CREST, PSE, and Sciences Po.

# 1 Introduction

To what extent is the income of individuals related to that of their parents? This question has seen renewed interest both in the general public and in academia as rising income inequality raised concerns about equality of opportunity. Examining this link is essential to understand whether children from different socio-economic backgrounds are afforded the same opportunities. It also matters for economic efficiency, as high persistence across generations may reflect an inefficient allocation of talents (so-called "Lost Einsteins"). Intergenerational persistence has now been estimated for a large number of countries, paving the way for insightful cross-country comparisons. Yet, much remains to be known for France, a country with relatively modest post-tax/transfers income inequality in international comparison and largely inexpensive higher education tuition fees.

The few existing studies for France only estimate the traditional intergenerational income elasticity (IGE), which captures the elasticity of child income with respect to parent income, and are based on small-sample surveys with self-reported incomes (Lefranc and Trannoy, 2005; Lefranc, 2018). Using a large sample combining census and tax return data, we estimate two additional measures of intergenerational mobility: (i) the rank-rank correlation (RRC) - the literature's new standard -, which corresponds to the correlation between child and parent income percentile ranks, and (ii) transition matrices, which capture finer mobility patterns along the parent income distribution. While previous studies on France used self-reported labor earnings, we focus on household-level income measures. They provide a better depiction of one's economic resources and allow the inclusion of children raised by single mothers. Integrating these improvements from the "new" intergenerational mobility literature enables us to conduct a detailed international comparison to rank France relative to other advanced economies.

In addition, we investigate the spatial variations in intergenerational mobility across the 96 metropolitan French departments. Such subnational analyses, pioneered by Chetty et al. (2014), help shed light on the mechanisms that may underlie income persistence across generations. Importantly, they highlight that national level estimates provide an incomplete assessment of a country's intergenerational mobility. We make use of the panel dimension of our data to describe the geographic mobility patterns of individuals and study the relationship between geographic mobility and intergenerational mobility. We investigate the separate roles of moving to a higher-income department from that of climbing the income ladder within departments, conditional on parent income rank.

Our analysis is conducted on almost 65,000 children born between 1972 and 1981, and observed in the Permanent Demographic Sample (EDP). This rich administrative dataset allows us to implement the contributions discussed above and to convincingly address concerns related to lifecycle and attenuation bias (Haider and Solon, 2006; Black and Devereux, 2011; Nybom and Stuhler, 2017). Since parents' incomes are not observed, we use the two-sample two-stage least squares estimation which consists in predicting parents' incomes using other parents drawn from the same population but for whom income is observed (Björklund and Jäntti, 1997). This method has been previously employed in the French context (Lefranc and Trannoy, 2005; Lefranc, 2018) as well as in many other countries (see Jerrim, Choi and Simancas (2016, Table A1)).

While studies typically use education and/or occupation to predict parent income, we make use of the richness of our data to also include detailed demographic characteristics (French nationality dummy, country of birth, household structure, and birth cohort), and characteristics of the municipality of residence (unemployment rate, share of single mothers, share of foreigners, population, and population density). Our results are largely insensitive to the set of predictors. Parent income is then defined as household-level predicted average annual pretax wage[1] over ages 35-45, and child income as pretax household income[2] averaged over the same age range between 2010 and 2016. These two income definitions represent the most comprehensive household-level income definitions possible for either generation.

**National Results.** Our main finding is that France exhibits relatively strong intergenerational income persistence compared to other developed countries. Our baseline estimate of the intergenerational elasticity in household income is 0.515, suggesting that on average, a 10% increase in parent income is associated with a 5.15%[3] increase in child income. Put differently, if one's parents earn 10% more than the average of parents' incomes, then one is expected to preserve about 50% of that relative advantage. Our father-son wage IGE, which can be compared to existing studies, equals 0.44, between Spain (0.40) and the United States (0.47), and far from Scandinavian countries (around 0.2) (Corak, 2016).

Moving to the rank-rank relationship, we find that the conditional expectation of child income percentile rank with respect to parent income percentile rank is linear throughout most of the parent income distribution, with steeper relationships at the tails. Our baseline estimate of the rank-rank correlation is 0.337, implying that a 10 percentile increase in parent income rank is associated, on average, with a 3.37 percentile increase in child income rank. This estimate is of similar magnitude to that

[1] Self-employment income is not observed and therefore not included in our parent income measure.

[2] Defined as the sum of labor earnings (wages + self-employment income), taxable capital income and predicted non-taxable capital income, unemployment insurance, retirement and alimony. Social benefits such as family allowances, social minima (e.g., RSA, disability benefits) and housing benefits are not included in this definition. See Section 3.3 for details.

[3] The exact expected change is equal to $(1.1^{0.515} - 1) \times 100 \approx 5.03\%$.

found for the United States (0.341; Chetty et al. (2014)), but markedly greater than existing estimates for other advanced economies such as Sweden (0.197; Heidrich (2017)), Australia (0.215; Deutscher and Mazumder (2020)) or Canada (0.242; Corak (2020)).[4]

Intergenerational persistence, as captured by the transition matrix, is strongest at the tails of the parent income distribution: 10.1% of children from the bottom 20% of the parent income distribution reach the top 20% as adults. This probability is almost 4 times greater for children born to parents in the top 20% (39.1%). In comparison, the probability for a child born to a family in the bottom 20% to reach the top 20% in adulthood is 7.5% in the United States (Chetty et al., 2014) and 12.3% in Australia (Deutscher and Mazumder, 2020).[5] Moreover, persistence at the top becomes stronger and stronger as we zoom in on the right tail of the parent income distribution.

We assess the robustness of our baseline results to a number of statistical biases. Foremost, we evaluate how sensitive they are to the lifecycle and attenuation biases by varying the ages at which child and parent incomes are measured as well as the number of parent income observations used. Our baseline results do not appear to under- nor over-estimate intergenerational mobility due to measuring child and/or parent incomes too early or too late in the lifecycle or because of averaging incomes over too few years. Moreover, we check whether using machine learning algorithms and varying the set of predictors influences our estimates. Slightly improved prediction from using flexible machine learning algorithms does not quantitatively alter our estimates. IGE estimates are overinflated when using only education as a predictor, while the RRC and transition matrices remain surprisingly stable regardless of the set of predictors used.

**Subnational Results.** We uncover substantial spatial variations in intergenerational mobility across departments, comparable to those observed across countries. We define individuals' location as their department of residence in 1990, when they are between 9 and 18 years old. Higher levels of mobility are typically found in the West of France, and lower levels in the North and South. While the IGEs range from 0.27 to 0.40 in departments in Brittany (West), they range from 0.46 to 0.71 in departments in Hauts-de-France (North). The distribution of department-level RRCs is tighter than that of IGEs, but displays very similar spatial patterns.

We also characterize departments' absolute upward mobility (AUM), defined as the expected income rank of children born to parents at the 25$^{\text{th}}$ percentile, which is obtained from the fitted values of the department-level rank-rank regression (Chetty et al., 2014). Absolute upward mobility ranges from the 34.4 in Pas-de-Calais (North) to 54.7 in Haute-Savoie (East). The Paris department stands out in terms of AUM

[4]See Table 1 and Appendix Figure E.6 for a comparison of RRC estimates.
[5]See Table 2 for a comparison of transition matrix estimates.

(52.3) but exhibits around average intergenerational persistence levels in terms of IGE (0.51) and RRC (0.31). The cross-department correlation between the IGE and RRC is only 0.65, and $-0.46$ with AUM. This highlights the importance of using a variety of intergenerational mobility measures to characterize a country's income persistence across generations (Mazumder and Deutscher, forthcoming).

Lastly, we conduct a descriptive analysis of the relationship between intergenerational income mobility and geographic mobility. We document important gains in expected income rank for movers, which are slightly decreasing in parent income rank. For children from families in the bottom decile, movers have an expected rank approximately 7 percentiles greater than stayers, while this difference is of roughly 3 percentiles for children from families in the top decile. We show that gains are partly attributable to movers locating in higher-income departments in adulthood relative to stayers, but also to movers reaching local ranks in their adulthood department that are further away from the rank of their parents in the childhood department. Destination departments are on average characterized by higher income levels than origin departments only at the tails of the parent income distribution. However, regardless of parent income rank, conditional on moving, the absolute upward mobility gains associated with moving to a higher-income department appear to be large and increasing with average income in the destination department. All these findings combine self-selection and causal effects, and we leave the disentangling of these two channels for future research.

The rest of the article is organized as follows. Section 2 describes the intergenerational income mobility measures we estimate and the main sources of bias they are subject to. The data, the parent income prediction procedure, and the sample and variable definitions are presented in Section 3. Section 4 reports our baseline estimates at the national level, while Section 5 assesses their robustness to various sources of bias. In Section 6, we investigate the spatial variations in intergenerational income mobility and describe the relationship between geographic and intergenerational mobility. Section 7 concludes.

# 2 Measuring Intergenerational Mobility

Intergenerational income mobility can be characterized using a variety of statistics.[6] In this section we (i) describe the statistics we employ, and (ii) discuss the two major

[6]See for example Corak (2020), where nine statistics of intergenerational mobility are put into perspective. More elaborate discussions on the properties of the different intergenerational mobility estimators can also be found in Black and Devereux (2011), Chetty et al. (2014), Nybom and Stuhler (2017), and Mazumder and Deutscher (forthcoming).

biases inherent to most intergenerational persistence estimators, namely lifecycle bias and attenuation bias.

## 2.1 Main Measures

Intergenerational persistence measures primarily aim to characterize the joint distribution of children and their parents' lifetime incomes with a parsimonious set of practical statistics. We summarize intergenerational persistence using the following statistics.

**Intergenerational Income Elasticity (IGE).** The traditional intergenerational income elasticity is obtained by regressing children's log lifetime income on their parents' log lifetime income. An IGE of 0.4 implies that a 10% increase in parent income is associated, on average, with a 4% increase in child income. Importantly, this estimator is sensitive to differences in inequality across generations. This can be seen in the following equation, where $y_p$ and $y_c$ are parent and child log lifetime incomes:

$$\text{IGE} = \frac{\text{Cov}(y_c, y_p)}{\text{Var}(y_p)} = \text{Corr}(y_c, y_p) \times \frac{\text{SD}(y_c)}{\text{SD}(y_p)}. \tag{1}$$

The empirical literature has highlighted that IGEs are particularly sensitive to lifecycle and attenuation biases, sample selection criteria, non-linearities along the parent income distribution, income definitions, and to the treatment of negative/zero incomes (Couch and Lillard, 1998; Chetty et al., 2014; Landersø and Heckman, 2017; Helsø, 2021).

**Rank-Rank Correlation (RRC).** The increasingly popular rank-rank correlation is obtained by regressing children's percentile rank in lifetime income on their parents' percentile rank in lifetime income. A RRC of 0.4 means that a 10 percentile increase in parent rank is associated, on average, with a 4 percentile increase in child rank. Unlike the IGE, the RRC is unaffected by inequality levels in either generation. This can be seen in the following equation, where $p_p$ and $p_c$ are parent and child percentile ranks in their respective lifetime income distributions:

$$\text{RRC} = \frac{\text{Cov}(p_c, p_p)}{\text{Var}(p_p)} = \text{Corr}(p_c, p_p) \times \frac{\text{SD}(p_c)}{\text{SD}(p_p)} = \text{Corr}(p_c, p_p). \tag{2}$$

Consequently, the greater the degree of inequality in the child generation relative to the parent generation, the greater the IGE relative to the RRC. In addition, the same RRC in two countries with large differences in inequality would hide that in one country the distance between ranks in monetary terms is actually much larger than in the other. The RRC owes its recent popularity to its robustness to specification variations, common biases, and treatment of negative/zero incomes (Dahl and DeLeire, 2008;

Chetty et al., 2014; Nybom and Stuhler, 2017).

**Transition Matrices.** To get a finer picture, one can use transition matrices, which report the probability of ending up in a given quantile as an adult conditional on coming from a family in a given quantile. Typically, they are reported by quintile and are of particular interest to seize non-linearities in children mobility across the parent income distribution.

## 2.2 Main Sources of Bias

The vast majority of currently available data sources do not cover the whole lifetime of children's and/or parents' incomes, leading researchers to approximate lifetime income based on shorter time spans. This data limitation generates the following two fundamental biases, which we extensively investigate in Section 5.

**Attenuation Bias.** A direct implication of relying on a limited number of income observations to approximate parent lifetime income is the attenuation bias arising from classical measurement error (Solon, 1992; Zimmerman, 1992). This leads to downward-biased estimates of intergenerational mobility. Mazumder (2005, 2016) and Nybom and Stuhler (2017) find that the attenuation bias can be very large for the IGE but affects the RRC only mildly, while O'Neill, Sweetman and Van de gaer (2007) show that it affects most the corner elements of the transition matrix. The common solution to lessen this bias is to average parent income over as many years as possible.

**Lifecycle Bias.** The second common bias relates to the age at which child and parent incomes are observed (Grawe, 2006; Haider and Solon, 2006). In particular, lifecycle bias arises in the presence of heterogeneous age-income profiles, which is observed empirically as high lifetime income individuals tend to experience steeper earnings profiles than low lifetime income individuals. As such, observing child or parent incomes either too early or too late in the lifetime is likely to bias intergenerational persistence estimates. The IGE is particularly sensitive to lifecycle bias, especially if incomes are measured before age 35, while it affects the RRC only moderately so long as incomes are measured at least in the late 20s/early 30s. Just as for the attenuation bias, the corner elements of the transition matrix are most sensitive to lifecycle bias (Chetty et al., 2014; Nybom and Stuhler, 2016, 2017).

# 3 Data

We use data from the Permanent Demographic Sample, which combines several administrative data sources on individuals born on the first four days of October.[7] We refer to individuals born on one of these days as *EDP individuals*. We describe below the most relevant details for each data source we use and provide additional technicalities in Appendix A.

**Civil Registers.** They contain information from birth certificates of EDP individuals and their children, including gender, date and place of birth, and parents' date and place of birth, nationality and occupation.

**1990 Census.** It contains socio-demographic information about EDP individuals and members of their household. Importantly, it reports parents' education level, occupation, and other demographic characteristics if EDP individuals live with their parents in 1990.

**All Employee Panel.** Since 1967, it gathers worker-year level information on all private and public sector employees in metropolitan France, except those in the agricultural sector.[8] Prior to 2001, only individuals born on an even year are covered.

**Tax Returns.** They provide tax information from 2010 to 2016[9] on individuals in dwellings where an EDP individual is known either from their income tax form or their main housing tax. Income variables are available both at the household level and at the individual level. An advantage of the information being gathered at the dwelling level is that household income is observed for all couples, regardless of whether they file their taxes jointly.

## 3.1 Parent Income Prediction

The measures of intergenerational mobility laid out in Section 2.1 cannot be estimated directly with our data since we do not observe parents' incomes. We therefore rely on the two-sample two-stage least squares (TSTSLS) strategy introduced by Björklund and Jäntti (1997).[10] It consists in predicting individuals' parents' incomes from a sam-

[7] See Robert-Bobée and Gualbert (2021) for a detailed description of the dataset. The EDP selection criterion has progressively widened to include individuals born on the first days of January, April, and July.

[8] See Appendix A for details on the coverage of the All Employee Panel.

[9] This corresponds to fiscal years 2011-2017.

[10] This method has been used in the French context by Lefranc and Trannoy (2005) and Lefranc (2018) and for many other countries where child and parent incomes cannot be observed simultaneously (Jer-

ple of other parents whose incomes are observed using a set of common observed characteristics. We refer to these other parents as *synthetic* parents.

Let $Z$ denote a set of characteristics observed both for parents and synthetic parents. Their log lifetime incomes $y$ can be expressed as:

$$y_i = \beta Z_i + \varepsilon_i. \tag{3}$$

We estimate this first-stage equation by OLS[11] on our sample of synthetic parents, and predict parents' log lifetime incomes using the resulting $\hat{\beta}$ as $\hat{y}_i = \hat{\beta} Z_i$. $Z$ includes (i) education (8 categories), (ii) 2-digit occupation (42 cat.), (iii) demographic characteristics (birth cohort, French nationality dummy, country of birth (6 cat.), and household structure (6 cat.)), and (iv) characteristics of the municipality of residence (unemployment rate, share of single mothers, share of foreigners, population, and population density).[12] For the geographic analysis, we drop the city characteristics to ensure they do not spuriously drive any spatial patterns. All characteristics are observed in the 1990 census. To reduce the potential for lifecycle and attenuation bias, synthetic parents' income is defined as average pretax wage between 35 and 45 with at least 2 income observations over this age range in the All Employee Panel. The model is estimated separately on synthetic mothers (adj. $R^2 = 0.37$ ) and fathers (adj. $R^2 = 0.36$).[13]

**Method Validity.** Despite the extensive use of the TSTSLS method for estimating the IGE, little is known about the consistency of this estimator. Using the Panel Study of Income Dynamics (PSID), Jerrim, Choi and Simancas (2016) compare estimates obtained using parents' observed incomes with those obtained by TSTSLS. They show that the sign and magnitude of the bias depends on the set of first-stage regressors and the number of synthetic parent income observations used.[14] Thus, we remain cautious about the exact magnitude of our IGEs. Such an exercise has not been performed for the RRC or transition matrices but we suspect they are significantly less sensitive to these issues as (i) they rely on income *ranks* instead of *actual* income, and (ii) they do not depend on the variance of parent income.

rim, Choi and Simancas, 2016, Table A1).

[11] In Appendix Section B.2 Figure B.2, we show that using more flexible and machine learning models does not alter our results.

[12] In Appendix Section B.2 Figure B.3, we show that our estimates are largely insensitive to the set of first-stage regressors, except for the IGE which is significantly larger when using only education in the first-stage.

[13] Appendix Figure E.1 shows that trimming the bottom (or the top) of the distribution does not improve the out-of-sample mean squared error, and Appendix Section B.4 Figure B.9b documents the influence of such trimming on the estimates.

[14] Acciari, Polo and Violante (forthcoming) perform a similar exercise using Italian administrative data from tax returns. They find a significant upward bias though use a limited set of predictors (father's age, province of birth and share of self-employment income).

**Inference.** Since we are in a two-stage setting, standard inference is inappropriate. Inoue and Solon (2010) derive an analytical formula for standard errors in the two-sample two-stage least squares setting. However, their method cannot be applied in our setting as we use a non-standard transformation of the first-stage outcome variables (i.e., predicted father income + predicted mother income) in the second stage. We thus compute bootstrap standard errors. Specifically, we draw one bootstrap sample for synthetic fathers and one for mothers separately. We then run the first-stage regression, and predict parent income on a bootstrap sample of children. We iterate this process 1,000 times. The bootstrap and naive standard errors are quantitatively similar for the baseline national and local results, and therefore, for computing time purposes, we report naive standard errors otherwise.

## 3.2 Sample Definitions

**Sample of Children.** It consists of EDP individuals who are (i) born between 1972 and 1981 in metropolitan France, (ii) observed with their parents in the 1990 census, (iii) whose parents are neither farmers nor in a liberal profession[15], and (iv) observed in the tax returns data at least once between 35 and 45 years old.[16] Restriction (i) is made to observe them with their parents in the 1990 census[17] and to have a reasonably large sample size for the subnational analysis. Restriction (ii) enables us to retrieve their parents' characteristics, and (iii) is due to the fact that farmers and liberal professions are not covered by the All Employee Panel from which we obtain synthetic parent income. Restriction (iv) aims to minimize lifecycle bias. The final sample contains 65,632 children.[18]

**Sample of Synthetic Parents.** It is constructed such that synthetic parents come from the same overarching population as actual parents. It therefore consists of EDP individuals who (i) had at least one child born between 1972 and 1981 in metropolitan France, (ii) are observed in the 1990 census, (iii) are neither farmers nor in a liberal profession in 1990, and (iv) have at least two pretax wage observations between 35 and 45 years old in the All Employee Panel.[19] As such our sample excludes individuals born

[15] 4.64% of EDP individuals satisfying (i) and (ii) have at least one parent who is a farmer and 2.1% have at least one parent who is in a liberal profession. As raised by Lefranc (2018), the fact that farmers tend to face relatively low incomes and a strong occupational inheritance (Lefranc, Pistolesi and Trannoy, 2009) makes the exclusion of farmers likely to bias intergenerational persistence downwards.

[16] 5.23% of EDP individuals satisfying (i) and (ii) are not observed in the tax returns data between 35 and 45 years old.

[17] See Appendix Figure E.2 for the position in the family in the 1990 census by child birth cohort.

[18] See Appendix Table F.1 for the sample size at each additional restriction.

[19] In Appendix Table F.2 we compare average characteristics of parents and synthetic parents. To ensure appropriate comparability of the two samples, no restriction on wage observations for synthetic parents or children is applied. Average characteristics are remarkably similar for most variables, even

in an odd year since they were not covered by the All Employee Panel prior to 2001. The final sample contains 31,423 synthetic parents.[20]

**Descriptive Statistics.** Appendix Table F.6 provides statistics on our sample of synthetic parents and children. On average, fathers are around 42 in 1990 and mothers 39. This assures that we predict income based on observable characteristics measured sufficiently late in their lifecycle.

## 3.3 Variable Definitions

The variables we use are constructed as follows. All incomes are expressed in 2015 euros, and are measured before taxes but after the deduction of employer- and employee-level payroll taxes.

**Parent Income.** We define the income of one parent as predicted average pretax wage over ages 35 to 45. This income is predicted according to the methodology described in Section 3.1. We then compute mean income at the household level (regardless of marital status) by taking the average of father and mother predicted incomes if the child is observed with both parents in the 1990 census[21], and income of the only parent otherwise. We refer to this income definition as parent household wage and use it as our main parent income measure. We also report results using father predicted income, which we refer to as father wage.

**Child Income.** Our main measure of child income, computed from the tax returns, corresponds to the sum of labor earnings (wages and self-employment income), taxable and imputed non-taxable capital income[22], unemployment insurance, retirement, and alimony, at the household level.[23] Just as for parents, a household is defined as individuals living in the same dwelling. To mitigate the potential for lifecycle bias, we average over 2010-2016 only for incomes declared when the individual is between 35

for 2-digit occupation (Appendix Table F.3), which confirms the assumption that actual and synthetic parents are random subsets of the same population.

[20] See Appendix Table F.4 for the sample size at each additional restriction.

[21] The parents observed in the same household as the child in the 1990 census do not necessarily correspond to the biological parents. Since we are interested in the relationship between the economic environment in which the child grew up and the child's own economic outcomes, the biological link is not relevant.

[22] Financial incomes not subject to any tax reporting are predicted by INSEE from a model estimated on the *Enquête Patrimoine*. In particular, they predict capital income for seven financial products (various tax-exempt savings accounts and life insurance) using household-level observed characteristics (income, age, family situation, ...). Excluding this income source from our child income definition does not affect the results.

[23] Social benefits such as family allowances, social minima (e.g., RSA, disability benefits) and housing benefits are not included in our main measure of child income.

and 45 years old.[24] We then divide by 2 for couples and complex households. We refer to this income definition as household income and use it as our main child income measure.

We also report results using the following alternative child income definitions: (i) household wage, which is equivalent to the parent household wage definition, (ii) individual income, which we define as the sum of all individual-level incomes: labor earnings (wages and self-employment income), unemployment benefits, retirement, and alimony, and (iii) individual wage.

**Income Definition Discussion.** Our preferred parent and child income definitions represent the most comprehensive household-level income definitions possible for either generation. Defining incomes at the household level is important in order to (i) better capture the economic conditions of individuals and their parents, (ii) allow the inclusion of children raised by single mothers, and (iii) enable the analysis of daughters, whose labor incomes alone may not be an appropriate measure of their economic outcomes. These income definitions are not identical but the results are qualitatively similar when using the same income definition, household wage, for both children and parents.

**Percentile Ranks.** We rank children within their birth cohort, and parents relative to other parents with children in the same birth cohort. Children with negative or zero incomes are assigned a rank equal to the ceiling of the percentage of such cases in their birth cohort divided by 2.[25] No parent has negative or zero predicted wage.

# 4 Results at the National Level

We start by analyzing intergenerational mobility at the national level. For our baseline results, we use data on children born on the first four days of October between 1972 and 1981 and measure parent income as household-level predicted average annual pretax wage over ages 35-45, and child income as pretax household income averaged over the same age range between 2010 and 2016. We include child birth cohort fixed effects in the log-log and rank-rank regressions.[26]

[24] Therefore, the 1981 birth cohort will only have at most one income observation, that for 2016 when they are 35; the 1980 cohort will have at most two income observations (2015 and 2016 when they are 35 and 36), etc. In that age range very few children still live in the same household as their parents (less than 5%).

[25] For example, if there are 3.65% of children with negative or zero incomes, they are assigned a rank of $\lceil 3.65/2 \rceil = 2$. Depending on the child income definition the percentage of children with negative or zero incomes varies. At most they represent 8% of the child sample, see Appendix Table F.5 for the exact figures.

[26] In practice, these fixed effects have virtually no influence on the coefficients of interest.

## 4.1 Intergenerational Income Elasticity (IGE)

Figure 1 panel A displays the conditional expectation of log child income with respect to log parent income. Children with negative or zero incomes are excluded. This is of minor importance when defining child income as household income as such cases are exceedingly rare.[27] The log-log CEF is pretty linear throughout the middle 80% of the parent income distribution, with some mild non-linearities at the tails.[28] This S-shaped relationship is also observed in the United States (e.g., Chetty et al. (2014)), Denmark (e.g., Helsø (2021)) or Sweden (e.g., Björklund, Roine and Waldenström (2012)). It implies that the elasticity is not constant over the whole parent income distribution, with smaller magnitudes at the tails, and is sensitive to the inclusion or exclusion of parents at the tails of their income distribution.[29]

Our baseline IGE estimate is 0.515, meaning that a 10% increase in parent income is associated, on average, with a 5.15%[30] increase in child income. Appendix Figure E.3 shows our estimates of the intergenerational income elasticity for every child and parent income definition, and for sons and daughters separately. Our father-son wage IGE estimate is relatively similar to existing ones for France despite important differences in methodology and data (see Appendix Table F.7). Intergenerational persistence estimates are larger for household income than for individual income or wage, which could be the result of assortative mating. IGEs are very similar when defining parent income as father wage, despite the fact that by construction, estimates based on father wage exclude children only observed with their mother in the 1990 census (10.29% of observations). The IGE is significantly lower for sons (0.467) than for daughters (0.563). This phenomenon is not systematic across countries, but is also observed in Germany (Bratberg et al., 2017) and the Netherlands (Carmichael et al., 2020), for instance.

## 4.2 Rank-Rank Correlation (RRC)

Figure 1 panel B plots the conditional expectation of child income rank with respect to parent income rank. It is relatively linear, with slight non-linearities at the tails as observed in many countries (Chetty et al., 2014; Bratberg et al., 2017; Helsø, 2021).

Our baseline estimate of the rank-rank correlation is 0.337, meaning that a 10 percentile increase in parent income rank is associated, on average, with a 3.37 percentile increase in child income rank. Appendix Figure E.4 shows our baseline estimates of the rank-rank correlation for every child and parent income definition, and for sons and

[27] We assess their influence for all child income definitions in Appendix Section B.4 Figure B.8.

[28] Appendix Figure B.3 shows that these non-linearities are not driven by the set of first-stage predictors.

[29] Appendix Section B.4 Figures B.9a and B.9c show how trimming the top and bottom of the parent/child income distribution influences our estimates.

[30] The exact expected change is actually equal to $(1.1^{0.515} - 1) \times 100 \approx 5.03\%$.

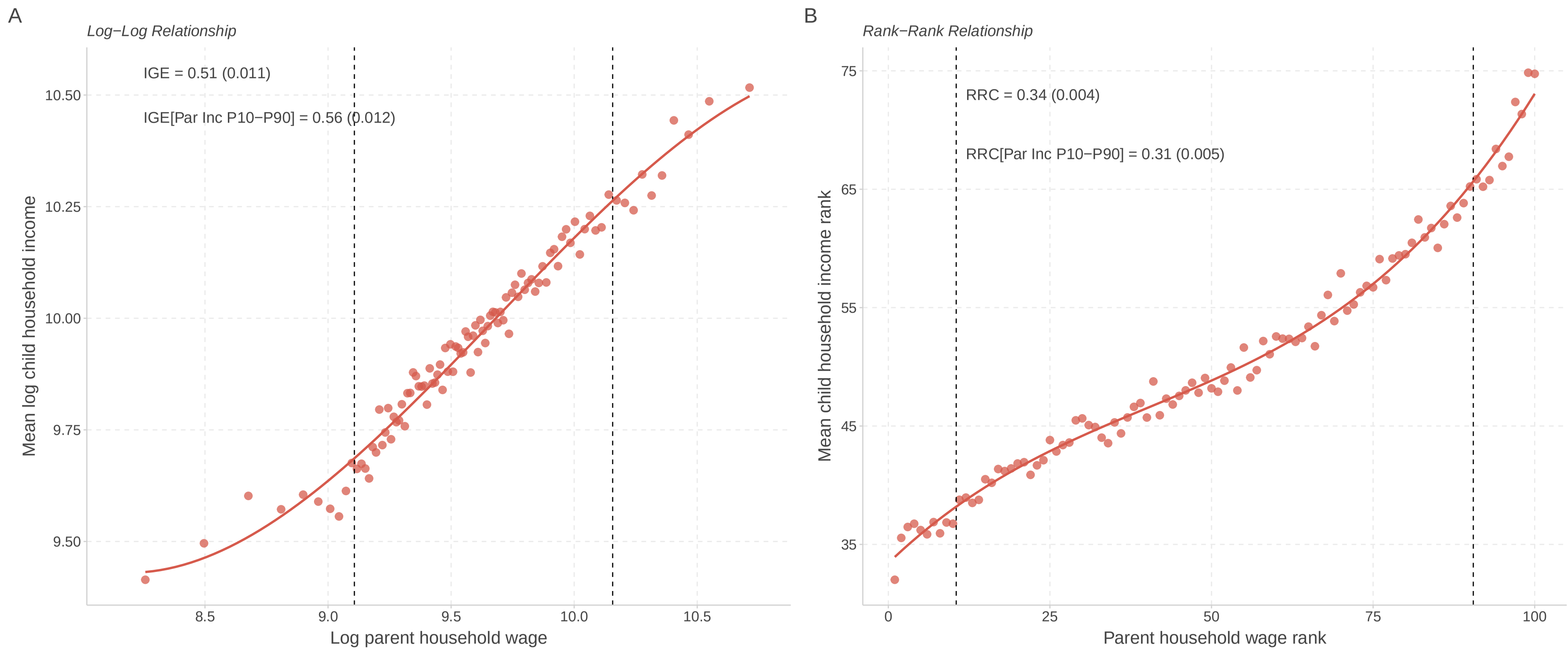



**Figure 1:** Conditional Expectation Functions for Log-Log and Rank-Rank Relationships

*Notes:* This figure presents non-parametric binned scatter plots of the relationship between log child income and log parent income (panel A), and child income rank and parent income rank (panel B). It is computed on the Permanent Demographic Sample, a dataset of individuals born on the first four days of October. The sample used is restricted to children born between 1972 and 1981. Child income is the mean of 2010–2016 household income (with age restricted to 35-45), divided by the number of household adults. Parent income is the sum of each parent predicted wage divided by the number of parents. Parent income is predicted separately for males and females using an OLS model including education (8 cat.), 2-digit occupation (42 cat.), demographic characteristics in 1990 (birth cohort, French nationality dummy, country of birth (6 categories), and household structure (6 cat.)) and characteristics of the municipality they lived in in 1990 (unemployment rate, share of single mothers, share of foreigners, population, and population density). These municipality characteristics are excluded for the geographic analysis. It is estimated on a sample of synthetic parents whose average wages at ages 35-45 (at least 2 income observations) is used as the dependent variable. Incomes are in 2015 euros. To construct panel A, children with negative or zero incomes are excluded (0.07% of the sample) and we bin parent incomes into 100 equal-sized bins and plot mean log child income versus mean log parent income within each bin. To construct panel B, children are ranked relative to other children in the same birth cohort while parents are ranked relative to other parents with children in the same birth cohort. Children with negative or zero incomes are assigned a rank equal to the ceiling of the percentage of such cases in their cohort divided by 2. We then plot mean child income rank versus parent income rank. The dashed lines represent the $10^{th}$ and $90^{th}$ percentiles of parents' income. We report coefficients and bootstrap standard errors (in parenthesis) obtained from OLS regressions of log child income on log parent income (panel A) and child income rank on parent income rank (panel B), both with child cohort fixed effects, on the microdata for the full sample and for parents between the $10^{th}$ and $90^{th}$ percentiles. The fitted line is a $3^{rd}$ order polynomial fit through the conditional expectations.

daughters separately. The estimates are slightly higher for daughters (0.351) than for sons (0.324), and are also slightly higher when defining parent income as household wage rather than as father wage. The estimates are significantly lower when defining child income as household wage or individual income and even lower when using individual wage, a pattern observed in other countries (Chetty et al., 2014; Deutscher and Mazumder, 2020; Landersø and Heckman, 2017), again possibly due to assortative mating.

To the best of our knowledge, this is the first time the RRC is estimated for France. In Table 1 we compare RRC estimates for countries for which estimates exist.[31] To enable comparability we only keep studies which pool sons and daughters together and define parent income as the sum or average of father and mother income.[32] Even though they are not directly comparable due to important differences in data and sample selection rules, we believe that it is a relevant exercise given the stability of the RRC to specification variations and common data limitations (e.g., observing child incomes only at relatively early ages (Nybom and Stuhler, 2017)).

This international comparison suggests that (i) France exhibits strong persistence across generations in international comparison, given that it is the country with the second highest available RRC estimate behind the United States, and (ii) there is less variation across countries in the rank-rank slope than in the intergenerational elasticity, which is coherent with the fact that the RRC is not influenced by changes in inequality across generations, and is less sensitive to sample restrictions.

## 4.3 Transition Matrices

The last measure of intergenerational income persistence we estimate is a quintile-by-quintile transition matrix, which documents the conditional probabilities of being in each income quintile as an adult given any parent income quintile. Figure 2 presents our baseline estimates of the transition matrix for France, along with available estimates for the United States (Chetty et al., 2014) and Australia (Deutscher and Mazumder, 2020). To the best of our knowledge, this is the first time transition matrices are estimated for France.[33]

We find that 10.1% of children born to parents in the bottom 20% reach the top 20% in their forties. This share is 7.5% in the United States and 12.3% in Australia.

[31] Appendix Figure E.6 provides a visual illustration of this Table.

[32] For most countries, child income is defined at the household or family level except in Chuard-Keller and Grassi (2021), Heidrich (2017) and Acciari, Polo and Violante (forthcoming) where it is at the individual level. For both parents and children, all the studies compiled use a comprehensive income definition.

[33] Alesina, Stantcheva and Teso (2018) estimated father-son wage transition probabilities from the bottom quintile only, using the TSTSLS methodology and data from the INSEE's *Formation et Qualification Professionnelle* for earlier cohorts (1963-1973).

| Country | RRC ↓ | # obs. | Data | Child Income Definition[1] | Child Cohort | Child Age or Year at Income Measurement | Parent Age or Year at Income Measurement | Source |
|---|---|---|---|---|---|---|---|---|
| Switzerland | 0.14 | 667,047 | Social Security Earnings Records | Average total pretax *individual* income | 1967-1982 | 32-34 | when child between 15-20 | Kalambaden and Martınez (2021, Table 3) |
| Switzerland | 0.14 | 923,262 | Social Security Earnings Records | Average total pretax *individual* income | 1967-1984 | 30-33 | when child between 15-20 | Chuard-Keller and Grassi (2021, Figure 1) |
| Spain | 0.195 | 1,492,107 | Atlas de Oportunidades | Total pretax *individual* income | 1980-1986 | 2016 | 1998 | Soria Espín (2022, Figure 1) |
| Sweden | 0.197 | 778,484 | SIMSAM database[2] | Average total pretax *individual* income | 1968-1976 | 32-34 | 34-50 | Heidrich (2017, Table 2) |
| Denmark | 0.203 | 157,543 | Danish register data | Average total pretax *family* income | 1980-1982 | 2011-2012 | 1996-2000 | Helsø (2021, Table 1) |
| Australia | 0.215 | 1,025,800 | Federal income tax returns | Average total pretax *family* income | 1978-1982 | 2011-2015 | 1991-2001 | Deutscher and Mazumder (2020, Table 2) |
| Sweden | 0.215 | 252,745 | 35% random sample from admin. data | Average total pretax *household* income | 1957-1964 | 1996-2007[3] | 1978-1980 | Bratberg et al. (2017, Table 3) |
| Norway | 0.223 | 324,870 | Full population admin. data | Average pretax *family* earnings | 1957-1964 | 1996-2006 | 1978-1980 | Bratberg et al. (2017, Table 3) |
| Canada | 0.242 | 2,115,150 | Intergenerational Income Data | Average total pretax *family* income | 1963-1970 | 2004-2008 | when child between 15-19 | Corak (2020, Table 5) |
| Germany | 0.245 | 1,128 | German Socio-Economic Panel | Average total pretax *household* income | 1957-1976 | 2001-2012 | 1984-1986 | Bratberg et al. (2017, Table 3) |
| Denmark | 0.253 | ≈ 410,000 | Danish register data | Average total pretax *family* income | 1973-1979 | 2010-2012 | when child between 7-15 | Landersø and Heckman (2017, Table A17) |
| Denmark | 0.257 | 205,625 | Full populations admin. data | Average total pretax *family* income | 1973-1975 | 2010-2012 | when child between 7-15 | Eriksen (2018, Table 3.2) |
| Italy | 0.30[4] | 1,719,483 | Electronic database of Personal Income returns | Average total pretax *individual* income | 1979-1983 | 2016-2018 | 1998-2000 | Acciari, Polo and Violante (forthcoming, p.28) |
| France | 0.337 | 64,572 | Permanent Demographic Sample | Parents: (Predicted) *household* wage; Children: average total pretax *household* income | 1972-1981 | 2010-2016 (between 35-45) | 35-45 | |
| United States | 0.341 | 9,867,736 | Federal income tax records, 1996-2012 | Average total pretax *family* income | 1980-1982 | 2011-2012 | 1996-2000 | Chetty et al. (2014, Table 1) |
| United States | 0.395 | 6,414 | NLSY79 | Average total pretax *family* income (self-reported) | 1957-1964 | 1996-2008[2] | 1978-1980 | Bratberg et al. (2017, Table 3) |

*Notes:*
[1] The parent income definition is always at the family level.
[2] Swedish Initiative for Research on Microdata in the Social and Medical Sciences.
[3] Only even years.
[4] This estimate corresponds to the one when adjusting for lifecycle bias, omission of taxpayers and tax evasion as reported on p.28. The baseline RRC estimate reported in Table 3 is 0.22.

**Table 1:** Rank-Rank Correlation in International Comparison

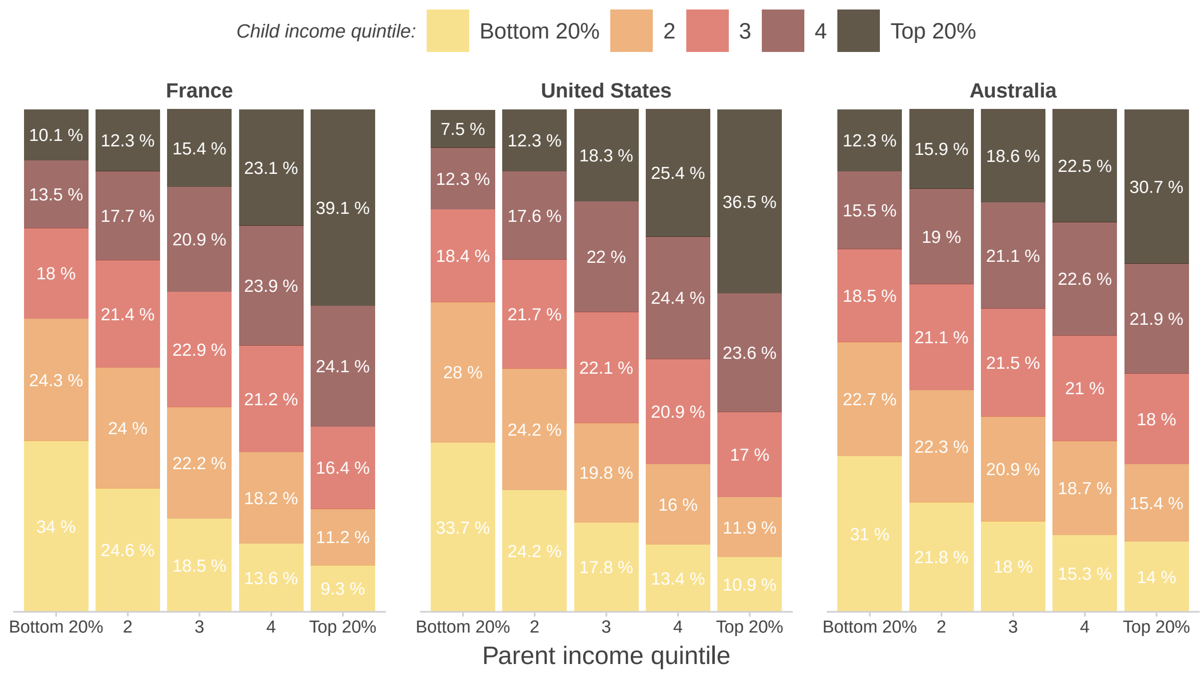



**Figure 2:** Baseline Quintile Transition Matrix for Different Countries

*Notes:* The first panel of this figure presents our baseline intergenerational transition matrix estimates. See Figure 1's notes for details on data, sample and income definitions. Each cell documents the share of children belonging to the quintile indicated by the color legend among children born to parents whose income falls in the quintile indicated on the x-axis. We present these estimates along with those put forward by Chetty et al. (2014) for the United States (second panel) and Deutscher and Mazumder (2020) for Australia (third panel).

In comparison, 34% remain in the bottom 20% of the income distribution. Regarding children born to the top 20%, 39.1% remain at the top, while only 9.3% move down to the bottom of the income distribution, much less than in Australia (14%). As a reference point, in a society where an individual's income is completely independent of parent income, the probability of being in any quintile given a parent quintile would by definition be 20%. We analyze persistence at the top of the parent income distribution in more detail in Appendix Section C.

Note that among the corner elements of the transition matrix, the estimates of mobility (i.e., P(Child Top 20% | Parent Bot. 20%) and P(Child Bot. 20% | Parent Top 20%)) are likely to be upper bounds, while estimates of persistence (i.e., P(Child Bot. 20% | Parent Bot. 20%) and P(Child Top 20% | Parent Top 20%)) are likely to be lower bounds. This is because the potential measurement error in parent rank prediction induced by TSTSLS can only go in one direction for the bottom and top quintiles. Parents in the bottom 20% necessarily have a true rank in the bottom 20% or above, but not below, as ranks take positive values by definition. Reasonably assuming that the probability of reaching the top 20% is increasing in parent income rank, our estimate of P(Child Top 20% | Parent Bot. 20%) is therefore likely to be an upper bound. The same reasoning can be applied to the other corner elements of the transition matrix.

| Country | P(Child Top 20% \| Parent Bot. 20%) ↓ | P(Child Bot. 20% \| Parent Bot. 20%) | P(Child Top 20% \| Parent Top 20%) | Source |
|---|---|---|---|---|
| United States | 7.5% | 33.7% | 36.5% | Chetty et al. (2014, Table 2) |
| Italy[1] | 8.6%[2] | 36.7% | 27.8% | Acciari, Polo and Violante (forthcoming) |
| France | 10.1% | 34% | 39.1% | |
| Denmark | 10.7% | 30.7% | 34.8% | Eriksen (2018, Figure 3.3) |
| Netherlands | 11.3% | 29.8% | 33.1% | Carmichael et al. (2020, Table 1) |
| Canada | 11.4% | 30.1% | 32.3% | Corak (2020, Table 6) |
| Switzerland | 11.9% | 23.7% | 30.3% | Chuard-Keller and Grassi (2021, Table 2) |
| Spain | 12.3% | 25.3% | 33.3% | Soria Espín (2022, Table A.5) |
| Australia | 12.3% | 31% | 30.7% | Deutscher and Mazumder (2020, Table 3) |
| Switzerland | 12.8% | 24.5% | 28.8% | Kalambaden and Martınez (2021, Table 5) |
| Sweden[3] | 15.7% | 26.3% | 34.5% | Heidrich (2017, Figure 10, Appendix B) |

*Notes:* See Table 1 for details about samples and income definitions used in each study.
[1] As the authors point out, this paper's baseline estimates are likely to overestimate upward mobility and underestimate persistence at the bottom and at the top because of lifecycle bias, the omission of taxpayers and tax evasion. The reported P(Top 20% | Bottom 20%) here corresponds to the estimate accounting as best as possible for these three sources of bias. For the other two measures, we report the estimates correcting for missing tax returns and tax evasion obtained from the authors.
[2] Obtained by multiplying the "Q1Q5" estimate found in the last column of Table 14 by the ratio of the two rows in Table 11, i.e., $0.100 \times 0.099/0.115$.
[3] Child incomes are measured relatively early in the lifecycle (32-34 years old), thus these estimates may suffer from lifecycle bias (i.e., overestimating upward mobility and underestimating persistence). By comparison, the father-son P(Child Top 20% | Parent Bot. 20%) estimate in Nybom and Stuhler (2017, Figure 1, Panel D) is essentially 10%, a much lower estimate of upward mobility.

**Table 2:** Transition Matrix in International Comparison

In Table 2 we compare conditional probabilities of interest with those found for other developed countries. In France income persistence across generations is particularly strong, both at the top and at the bottom. While France does better than the United States when it comes to upward mobility from the bottom quintile (10.1% vs. 7.5%), a point we discuss in Section 4.4, it fares significantly worse than countries such as Canada (11.4%), Switzerland (11.9%) or Australia (12.3%). It also displays one of the strongest persistence at the bottom and at the top of the income distribution.

## 4.4 Discussion of Baseline Results

**International Comparison.** Our findings confirm the conventional wisdom that France exhibits strong income persistence across generations relative to many OECD countries (OECD, 2018). This is true not only with respect to the IGE, which has been the main focus for cross-country comparisons in the literature (e.g., see Corak (2016)), but also for the RRC, and in terms of transition matrices. This raises the question of the underlying mechanisms. Indeed, one apparent puzzle is that various studies have found positive effects of government spending on intergenerational mobility (Mayer and Lopoo, 2008; Huang, Huang and Shui, 2021). Yet, despite significant government spending, France displays relatively little intergenerational mobility.

However, though the IGE and RRC estimates are very similar for France and the United States, the two countries differ in terms of the probability of reaching the top 20% conditional on having parents in the bottom 20%. Given the large dissimilarities in their higher education systems, part of the explanation could stem from differences in access to, and graduation from, higher education along the parent income distribution.

**Access to and Graduation from Higher Education.** Using the yearly census surveys available since 2004 in the EDP, we can observe children's last obtained diploma when they are between 23 and 45.[34] Figure 3 compares higher education *graduation rates* in France with *enrollment rates* in the United States[35] (Chetty et al., 2020) by parent income rank. Graduation rates in France are lower than enrollment rates in the United States, which is expected considering that a sizable share of students who enroll in higher education eventually drops out. While the relationship between parent income rank and enrollment is linear in the United States, obtaining a higher education degree appears to be a convex function of parent income rank in France. In particular, it is flatter at the bottom of the distribution.[36] This convex relationship is all the more striking since children from low-income families are probably more likely to drop out from higher education, and therefore not earn a higher education degree.

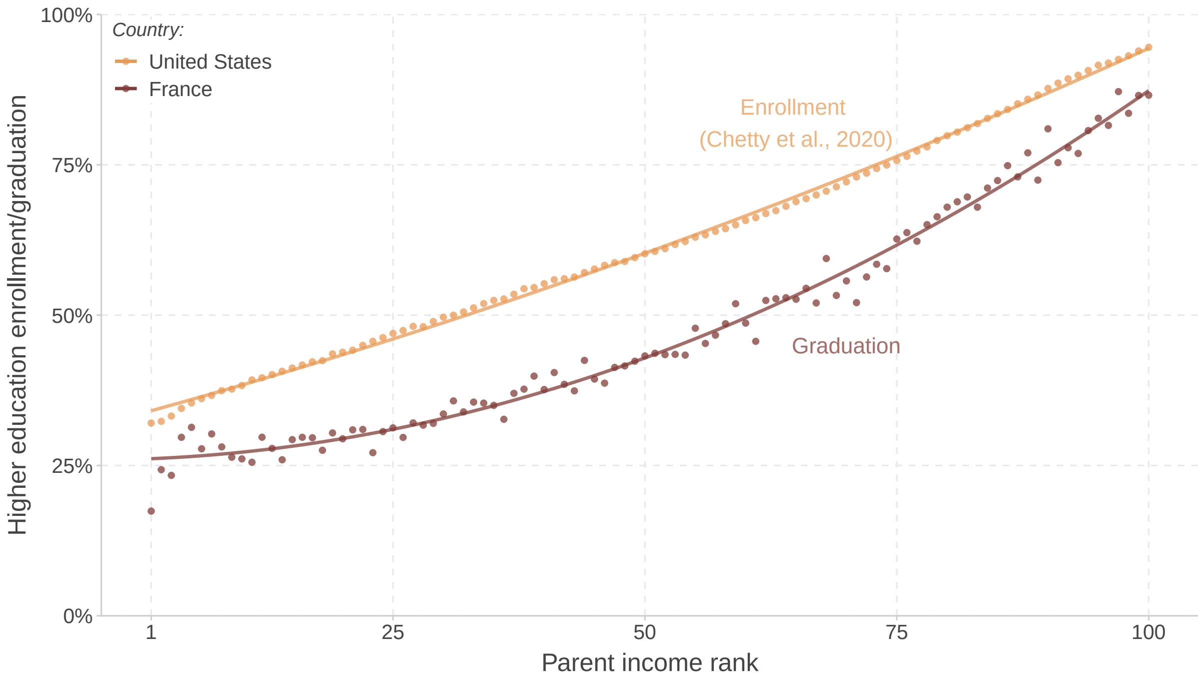



**Figure 3:** Graduation From/Enrollment In Higher Education by Parent Income

*Notes:* This figure presents higher education graduation in France vs. enrollment rates in the United States (Chetty et al., 2020) by parent income rank. See Figure 1's notes for details on data, sample and income definitions.

This comparison does not allow us to assess directly whether higher education

[34] We observe this information for 86.29% of the sample. The share of missing values is pretty well uniformly distributed along the parent income rank distribution.

[35] Specifically, enrollment is defined as attending college at least at some point between ages 18-21.

[36] Appendix Figure E.7 documents the graduation rate for each cell of the quintile-by-quintile transition matrix. It shows that the convexity in the relationship between family background and graduation rate holds within child income quintile.

may explain the gap in upward mobility between France and the United States, since the relationship between college completion and parent income rank for the latter is not available. Using a French survey of roughly 6,000 18-24 year olds, Bonneau and Grobon (2022) find that enrollment rates in higher education by parent income rank are very similar in France compared to the United States. Therefore, if higher education were to explain part of the upward mobility gap observed between the two countries, it must necessarily be trough differences in dropouts rates and/or heterogeneous returns to higher education along the parent income distribution.

# 5 Robustness of Baseline Results

## 5.1 Lifecycle and Attenuation Bias

As discussed in Section 2.2, the existing literature has highlighted two statistical biases that may affect our baseline estimates: lifecycle and attenuation bias. The former relates to heterogeneous lifecycle earnings profiles among parents and children, while the latter refers to classical measurement error in parent income. We therefore assess how our estimates vary with the age at which child and parent incomes are measured, and with the number of synthetic parent income observations used.

**Child Lifecycle Bias.** Figure 4 presents our estimates of intergenerational income mobility when varying the age at which child income is measured. In addition to household income from the tax returns data, we exploit the longer time series wage data provided by the All Employee Panel. Each point represents the estimate of the measure of intergenerational income mobility when measuring child income at a given age.[37] For the transition matrix, we only present the analysis for the conditional probability of being in the top or bottom 20% for children born to parents in the top or bottom 20%.

The broad pattern that emerges in Figure 4 panels A and B is that the estimated IGE and RRC increase sharply when child incomes are measured early in the lifecycle and stabilize roughly when child income is measured around 30 years old. The wage IGE (RRC) measured at age 25 is equal to 0.18 (0.12) while it is 0.40 (0.26) at age 35, more than a doubling in magnitude.[38] For household income there appears to be a slight

[37] By construction, each age estimate is obtained from a different sample since we only measure child incomes in the tax returns data between 2010 and 2016, and in the All Employee Panel from 1967 to 2015 (though only for individuals born in even years before 2001).

[38] A recent INSEE report by Abbas and Sicsic (2022) provides intergenerational mobility estimates for France using a subsample of individuals born in 1990 (i) who are still claimed as dependent in their parents' tax return at age 20, (ii) whose parents' income can be observed around age 50, and (iii) whose individual income is observed at age 28 in their own tax return. They compare their results to ours and despite different sample definitions, when using the same income definition and measuring child income at the same age (i.e., 28), they find very similar results.

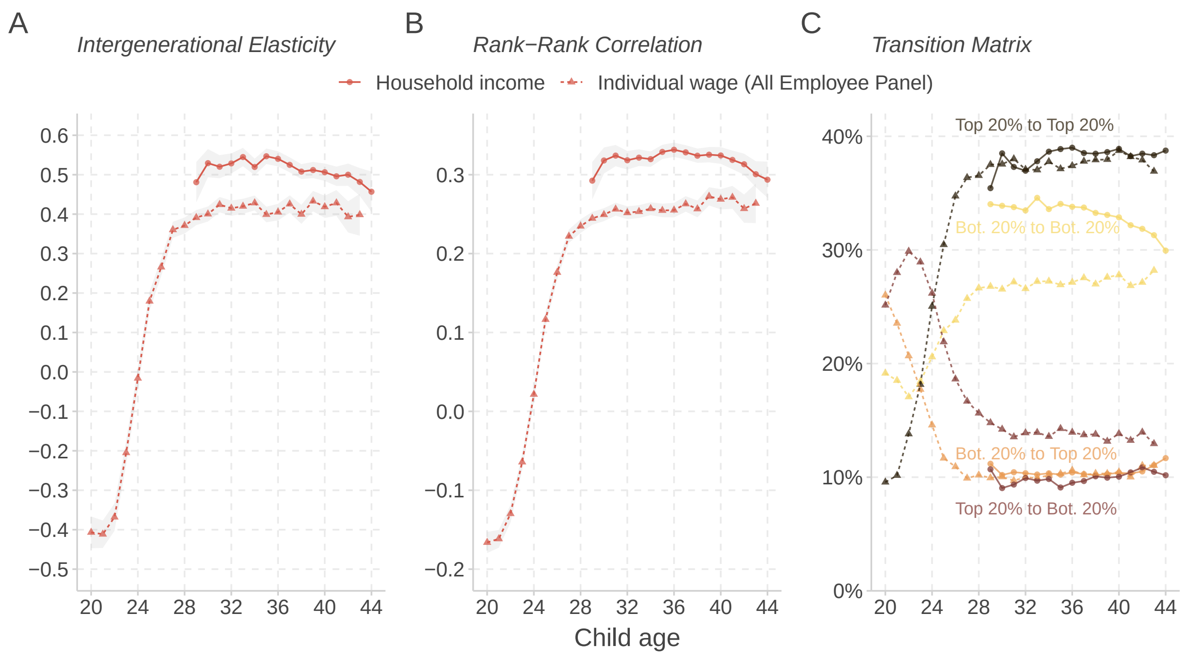



**Figure 4:** Child Lifecycle Bias

*Notes:* This figure assesses the robustness of our baseline intergenerational income mobility estimates to changes in the age at which child income is measured. Shaded areas represent the 95% confidence intervals. See Figure 1's notes for details on data, sample and income definitions.

decline in the estimates when children are in their forties. This appears to mostly reflect changes in the underlying cohort sample rather than a real decrease in the estimate.[39]

The results for the transition matrix in Figure 4 panel C suggest our baseline estimates are quite close to the estimates obtained when child income is measured at any age between 29 and 44, except for persistence in the bottom 20% which declines after age 36. The estimates using the All Employee Panel confirm that when measuring child incomes too early in the lifetime, the secondary diagonal elements of the transition matrix (remaining in the same income quintile as one's parents) would be severely underestimated while "big transitions" (from bottom to top and from top to bottom) would be severely overestimated.

Overall, we do not find persuasive evidence that the IGE or the RRC varies importantly with the age at which child income is measured so long as it is measured at least in their early thirties. This does not imply that there is no remaining lifecycle bias, as highlighted by Nybom and Stuhler (2016), it is simply suggestive that our baseline results do not measure child incomes too early in their lifecycle, where lifecycle bias would be larger.

**Parent Lifecycle Bias.** We assess the sensitivity of our baseline estimates to varying the

[39]In Appendix Section B.3 Figure B.4 we reproduce the All Employee Panel estimates keeping the sample of children constant.

age at which parent income is measured. Since we predict parent income rather than observe it, we vary the age at which synthetic parent income is measured in the first-stage regression. Specifically, we run the first-stage regression (3) defining synthetic parent income at a given age between 25 and 60 years old. Figure 5 shows how our estimates of intergenerational mobility vary with the age at which parent income is predicted. The relationship between age at which parent income is measured and both the IGE (panel A) and the RRC (panel B) is concave, strongly increasing between 25 and the late thirties and then stabilizing until the mid to late fifties. Relative to our baseline estimate, it does not appear that our choice of measuring synthetic parent income as the average between 35 and 45 years old (with at least 2 income observations) is either too early or too late in the lifecycle.

This mismeasurement of parent income also affects estimates of transition probabilities (Figure 5 panel C). Relative to our baseline results, measuring parent income at age 25 underestimates the likelihood of remaining in the bottom 20% or top 20%, and overestimates the probability of moving upwards (bottom 20% to top 20%) or downwards (top 20% to bottom 20%). The estimates stabilize once again when parent income is measured after age 35.[40]

[40]In Appendix Section B.3 Figure B.5 we study how our measures of intergenerational persistence vary with the age at which child and synthetic parent income is measured jointly.

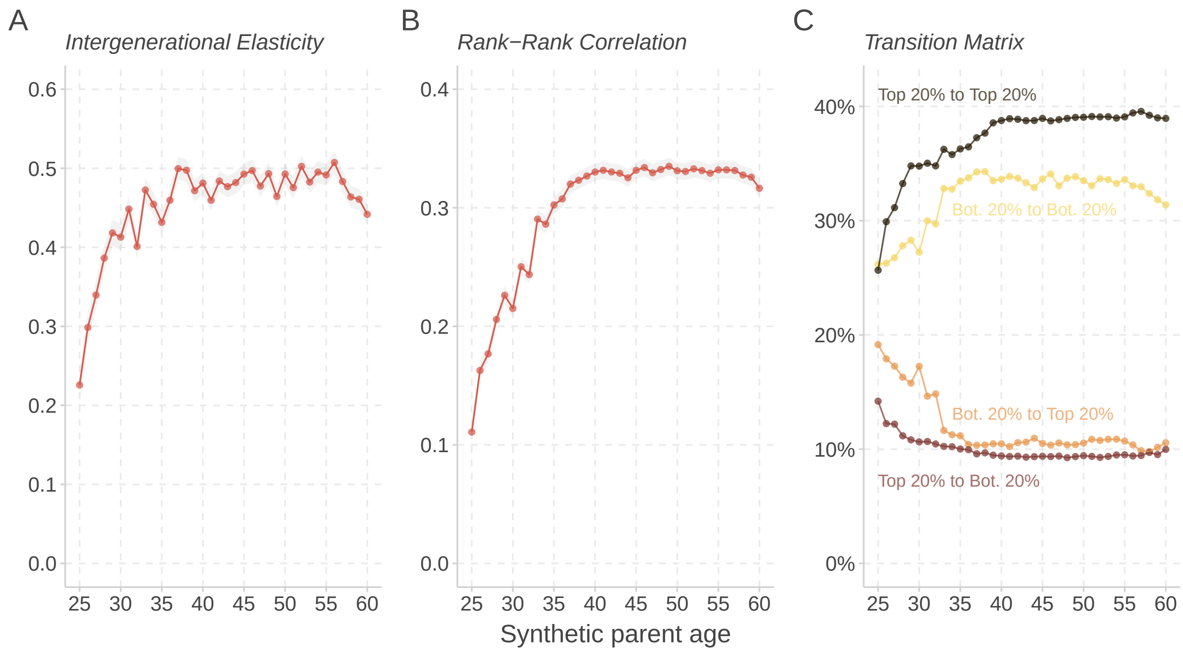



**Figure 5:** Parent Lifecycle Bias

*Notes:* This figure assesses the robustness of our baseline intergenerational income mobility estimates to changes in the age at which synthetic parent income is measured. Shaded areas represent the 95% confidence intervals. See Figure 1's notes for details on data, sample and income definitions.

**Attenuation Bias.**

We evaluate the extent to which our baseline estimates are sensitive to the number of observations used to compute parent lifetime income. The main source of attenuation bias comes from measurement error in parent income.[41]

Figure 6 plots estimates of our persistence measures varying the number of synthetic parent income observations used in the first-stage regression from 1 to 11. To control for the potential effect of lifecycle bias we center the age at which synthetic parent income is measured at 40 years old. In other words, one income observation corresponds to income at age 40, two income observations corresponds to average income at ages 39 and 41, three income observations to average income between 39 and 41, and so on. Therefore, 11 income observations corresponds to the average between 35 and 45 years old. The sample of synthetic parents over which income is predicted varies for each estimate depending on how many synthetic parents had incomes observed each year in the required age range.[42] We report results both for parent household wage and father wage.

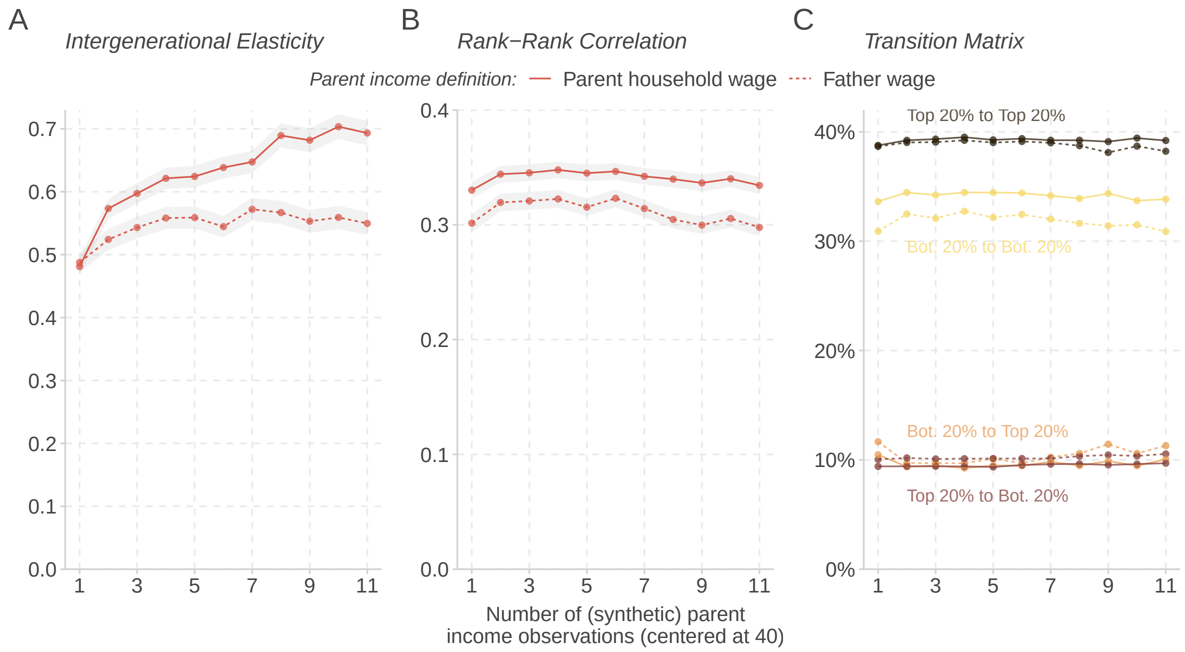



**Figure 6:** Attenuation Bias

*Notes:* This figure assesses the robustness of our baseline intergenerational income mobility estimates to the number of income observations used to predict parent income. While varying the number of parent income observations, we center the age range at 40 to control for lifecycle bias. Shaded areas represent the 95% confidence interval. See Figure 1's notes for details on data, sample and income definitions.

[41]We also check in Appendix Section B.3 Figure B.6 the sensitivity of intergenerational mobility to the number of child income observations and confirm that it only plays a very minor role.

[42]In Appendix Section B.3 Figure B.7 we reproduce the results keeping the sample of synthetic parents constant. It does not change the conclusions.

These results suggest that attenuation bias might affect our baseline IGE (panel A) but not our other estimates of intergenerational mobility. Indeed when defining parent income at the household level, the IGE increases from just below 0.5 when using only one income observation to around 0.7 when averaging over 11 income observations (i.e., between 35 and 45). It is important to highlight that almost all of this change is driven by how mothers' incomes are predicted.[43] Indeed when looking at the father-child IGE, the estimate does not increase so markedly and stabilizes around 2 or 3 income observations, consistent with the idea that the two-stage procedure employed drastically shrinks the transitory component of annual income, and in large contrast with what is typically found when parent income is actually observed (Mazumder, 2005). Indeed, since we are already predicting parent income based on observable characteristics, and thus in a sense reducing year-to-year income volatility, averaging over more years does not affect the estimate much.

The rank-based measures, whether the RRC (Figure 6 panel B) or the transition matrix cells (Figure 6 panel C), are remarkably unaltered by increasing the number of income observations over which synthetic parent income is averaged. In the context of TSTSLS estimation, this appears to be a strength of rank-based measures since it suggests that in cases where parent income is not observed, predicting it using only one synthetic parent income observation is likely to provide sufficiently accurate estimates.

# 6 Geographic Analysis

## 6.1 Heterogeneity Across Departments

A first step in understanding the sources of intergenerational mobility in France is to investigate where persistence is highest and lowest. We study the geographic variations of intergenerational mobility at the department level. Departments divide metropolitan France into 96 territories.[44] Departments have the advantage of covering the whole of metropolitan France, and their borders have not changed over the study period. In addition, considering a finer geographic unit such as commuting zones would imply dropping a sizable amount of areas due to insufficient sample size.

Children are assigned to their department of residence in 1990, when they were between 9 and 18 years old. This is the best proxy we have for the department they

[43] How one interprets the results based on parent household wage depends on one's prior as to how to best predict mothers' incomes. Our view is that predicting mothers' incomes only on the subsample of synthetic mothers with observed wages in all years between 35 and 45 years old might bias the underlying sample considering the uneven labor force participation of women at the time. We believe our choice of restricting our sample of synthetic parents to those with at least two income observations between ages 35 and 45 is reasonable.

[44] For practical reasons, we treat Corsica as a single department. Appendix Figure E.8 shows a map of French departments.

grew up in. To ensure our estimates are sufficiently reliable, we focus on the 85 departments with at least 200 observations.[45] Hereinafter we use parent income predicted without municipality characteristics in the first stage. This is to make sure that they do not spuriously drive any spatial patterns.[46] Moreover, we find that spatial variations in intergenerational mobility are not driven by differences in prediction accuracy of the first-stage across departments.[47] Individuals are still ranked within the national income distribution.

The statistics we use at the subnational level are (i) the IGE, (ii) the RRC, and (iii) the expected income rank for individuals whose parents locate at the 25th percentile, which we refer to as *absolute upward mobility* (AUM) following Chetty et al. (2014).[48] Denoting $p_{c,d}$ the percentile income rank of children observed in department $d$ during childhood, and $p_{p,d}$ the percentile income rank of their parents, local RRCs are obtained from the following OLS regression:

$$p_{c,d} = \alpha_d + RRC_d \times p_{p,d} + \varepsilon_d \tag{4}$$

The expected income rank for individuals whose parents locate at the 25th percentile then writes:

$$\text{AUM} := \mathbb{E}[p_{c,d} \mid p_{p,d} = 25] = \hat{\alpha_d} + R\hat{R}C_d \times 25 \tag{5}$$

Appendix Figure E.9 graphically illustrates how this intergenerational mobility measure is computed for the Nord department, the most populated one in 1990. The conditional expectation functions for the most populated departments are available in Appendix Figures E.10 and E.11. Even at the department level, it appears that the rank-rank relationship is well approximated by a linear function.

Figure 7 depicts department-level intergenerational mobility as captured by the three estimators mentioned above, and reveals substantial variations though not necessarily statistically significant, likely due to lack of statistical power.[49] The distribution of department-level RRCs ranges from 0.20 to 0.44 and is tighter than that of IGEs,

[45] The number of observations per department is reported in Appendix Table F.9.

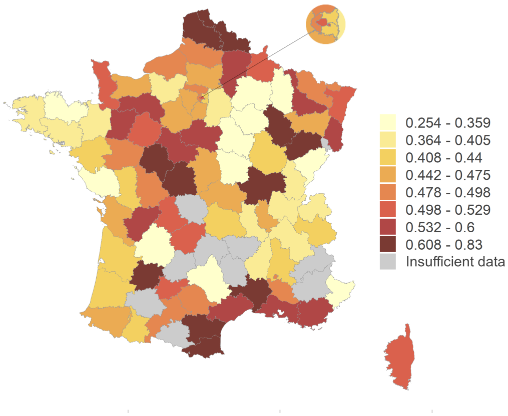
[46] The removal of municipality characteristics from the first stage does not alter our national estimates (see Appendix Figure B.3) nor the first-stage adjusted $R^2$ (from 0.46 to 0.45, when fully interacting predictors with gender). Moreover, the cross-department correlation with and without city characteristics is above 0.97 for all three intergenerational mobility measures (IGE, RRC, AUM).

[47] Indeed, as shown in Appendix Table F.8, the department-level mean-squared errors of the first-stage predictions are not significantly related with department-level intergenerational mobility estimates.

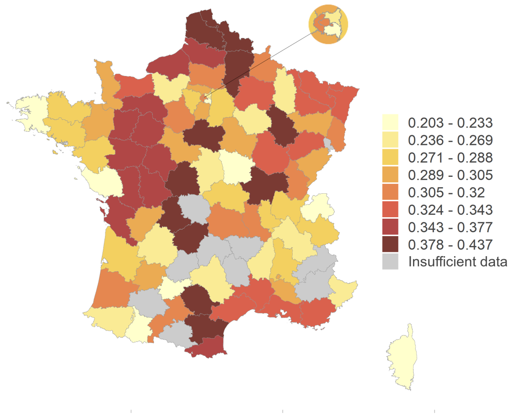
[48] We favor absolute upward mobility over specific cells of the transition matrix because of the size of our department samples. Indeed, while absolute upward mobility is estimated using all the observations in a given department, any cell of the quintile transition matrix is by construction estimated using only a fifth of these observations.

[49] Department-level estimates are reported in Appendix Table F.9. Department-level intergenerational elasticities and rank-rank correlations are represented graphically with their confidence intervals in Appendix Figures E.12 and E.13.

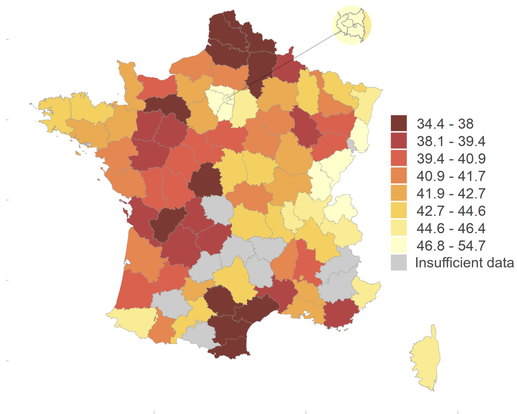
which ranges from 0.25 to 0.83. Both vary across departments just as much as they vary across countries. The range of our estimates of absolute upward mobility, from rank 34 to rank 55, is almost identical to that observed in Italy using a comparable geographic unit (from 35 to 57 (Acciari, Polo and Violante, forthcoming)).[50]




**(a)** Intergenerational Elasticity




**(b)** Rank-Rank Correlation




**(c)** Absolute Upward Mobility

**Figure 7:** Spatial Variations in Intergenerational Mobility

*Notes:* This figure presents department-level estimates of our intergenerational mobility measures. To compute local estimates, individuals are assigned to their department of residence in 1990, when they were between 9 and 18 years old. Departments with less than 200 observations are considered as having insufficient data. See Figure 1's notes for details on data, sample and income definitions.

Intergenerational persistence is particularly high in the North and in the South of France, and relatively low in the West. For instance, the IGEs range from 0.27 to 0.40 in departments in Brittany (West), from 0.46 to 0.71 in departments in Hauts-de-France (North), and from 0.60 to 0.73 in the former region of Languedoc-Roussillon (South). This pattern is observed not only in terms of relative mobility (IGE and RRC), but also in terms of absolute upward mobility. Indeed, while children with modest socio-economic backgrounds have relatively high expected income ranks in Brittany (AUM

[50]Using a finer geographic unit, Chetty et al. (2014) find that 80% of AUMs are between 37 and 52 in the United States.

$\in$ (42.5; 44.6)), they tend to remain lower in the income distribution in Hauts-de-France (AUM $\in$ (34.4; 41.5)) and Languedoc-Roussillon (AUM $\in$ (34.6; 38.3)). These spatial variations match quite closely those of the unemployment rate. In Appendix Section D we present a more detailed correlation analysis with local characteristics.

However, a high relative mobility is not systematically associated with a high absolute upward mobility. For instance, such a discrepancy is observed for the city-department of Paris, the third highest department in terms of AUM, but where intergenerational mobility levels in terms of IGE and RRC are close to the department-level average. The conditional expectation functions in Appendix Figure E.11 provide an explanation to this idiosyncrasy. They reveal that the Parisian CEF is both shifted upwards relative to other large departments, and flatter at the lower end of the parent income distribution. The combination of these two features results in relatively good prospects for children whose parents locate at the 25$^{\text{th}}$ percentile without implying particularly high relative mobility. The cross-department correlation between the IGE and RRC is 0.65, and is $-0.46$ with AUM, which highlights the importance of using a variety of intergenerational mobility measures to characterize a country's income persistence across generations (Mazumder and Deutscher, forthcoming).[51]

## 6.2 Geographic Mobility

Few studies have explored the relationship between geographic mobility and intergenerational mobility.[52] We consider individuals as geographically mobile if their adulthood department of residence is different from their childhood department of residence. The childhood department of residence is observed in the 1990 census, when individuals were aged from 9 to 18 years old. The adulthood department of residence is the one indicated on individuals' tax return. If the individual has lived in several departments over 2010-2016, we consider the most common department of residence. In case of ties, we consider the most recent of the most common departments. According to this definition, 41.8% of individuals are geographically mobile. This share is relatively homogeneous across males (41.4%) and females (42.3%). The percentage of movers by parent household wage rank is presented in Appendix Figure E.14.

**Intergenerational Mobility Gains from Geographic Mobility.** Figure 8 shows the conditional expectation of child household income rank with respect to parent household wage rank for movers and stayers. The CEF is slightly flatter for movers than for stayers, and importantly, movers have systematically higher expected income ranks

[51] Appendix Table F.10 reports the correlation between each intergenerational mobility measure for the three income definitions we use.

[52] Existing studies rather exploit geographic mobility to estimate the causal impact of location on upward mobility (Chetty and Hendren, 2018; Laliberté, 2021).


**Figure 8:** Intergenerational Mobility and Geographic Mobility

*Notes:* This figure represents the conditional expectation of child household income rank with respect to parent household wage rank separately for individuals whose adulthood department of residence is different or not from their childhood department of residence. Percentile ranks are computed according to the national income distribution, which implies that the share of movers and stayers is not constant throughout the parent income distribution. The childhood department of residence is observed in the 1990 census, when individuals were aged from 9 to 18 years old. The adulthood department of residence is the one indicated on individuals' tax return. If the individual has lived in several departments over 2010-2016, we consider the most represented department of residence. In case of ties, we consider the most recent of the most represented departments. See Figure 1's notes for details on data, sample and income definitions.

than stayers throughout the parent household wage rank distribution. The difference between the two CEFs is slightly decreasing in parent income and is particularly pronounced at the bottom of the distribution. This difference is the result of the combination of individuals self-selecting into migration and the causal effect of moving.

To characterize the relationship between intergenerational and geographic mobility, we estimate the following regression model:

$$p_{c,i} = \alpha + \beta p_{p,i} + \gamma \text{Mover}_i + \delta p_{p,i} \times \text{Mover}_i + X_i'\lambda + \varepsilon_i, \tag{6}$$

where $p_{c,i}$ is the household income rank of individual $i$, $p_{p,i}$ is individual $i$'s parents' household wage rank, $\text{Mover}_i$ is a binary variable taking the value 1 if individual $i$ lives in a different department from the one they grew up in and 0 otherwise, and $X_i$ is a set of control variables. Table 3 reports the corresponding regression results.

Column (1) shows the estimates from equation (6). Living in a different department from one's childhood department is associated, on average, with a $\mathbb{E}[\hat{\gamma} + \hat{\delta} p_{p,i}] = 5.95$

| | *Dependent variable:* Child household income rank | | | | |
|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) |
| Parent income rank ($\hat{\beta}$) | 0.311*** | 0.311*** | 0.299*** | 0.187*** | 0.137*** |
| | (0.005) | (0.005) | (0.005) | (0.007) | (0.014) |
| Mover ($\hat{\gamma}$) | 6.152*** | 6.183*** | 5.787*** | 6.008*** | 5.971*** |
| | (0.452) | (0.452) | (0.452) | (0.452) | (0.452) |
| Parent income rank × Mover ($\hat{\delta}$) | −0.004 | −0.004 | −0.003 | −0.019** | −0.020*** |
| | (0.008) | (0.008) | (0.008) | (0.008) | (0.008) |
| Constant | 32.332*** | 31.897*** | 35.663*** | 30.109*** | 32.502*** |
| | (0.407) | (0.422) | (1.569) | (1.844) | (1.922) |
| Birth cohort | ✓ | ✓ | ✓ | ✓ | ✓ |
| Gender | | ✓ | ✓ | ✓ | ✓ |
| Department FE | | | ✓ | ✓ | ✓ |
| Parents' education | | | | ✓ | ✓ |
| Parents' 2-digit occupation | | | | | ✓ |
| $\mathbb{E}[\hat{\gamma} + \hat{\delta} p_p] = \hat{\gamma} + \hat{\delta} \times 50.5$ | 5.95 | 5.98 | 5.64 | 5.05 | 4.96 |
| $\mathbb{E}[\hat{\gamma} + \hat{\delta} p_p \mid p_p = 25]$ | 6.05 | 6.08 | 5.71 | 5.53 | 5.47 |
| $\mathbb{E}[\hat{\gamma} + \hat{\delta} p_p \mid p_p = 75]$ | 5.85 | 5.88 | 5.56 | 4.58 | 4.47 |
| Observations | 64,572 | 64,572 | 64,572 | 64,572 | 64,572 |
| Adjusted $R^2$ | 0.118 | 0.118 | 0.128 | 0.142 | 0.148 |

*Notes:* This table provides the estimates from regression child household income rank on their parents' income rank, a dummy variable indicating whether the individual is a mover, and the interaction between these two variables. Columns (2) to (5) progressively include control variables. See Figure 8 for details on variable and sample definitions. *p<0.1; **p<0.05; ***p<0.01.

**Table 3:** Intergenerational & Geographic Mobility

percentile rank increase in the national household income distribution. The rank-rank slope starts to significantly differ between movers and stayers when controlling for parents education (col. (4)). In the full control specification (col. (5)), the difference in expected income rank between movers and stayers is decreasing in parent income (5.47 at the 25$^{th}$ percentile and 4.47 at the 75$^{th}$ percentile). The expected rank increase associated with geographic mobility decreases slightly with these additional controls but the rank difference between movers and stayers from the bottom of the parent income distribution remains particularly stable.

**The Role of Mobility Toward Richer Departments at the Aggregate Level.** There are several potential reasons for the better intergenerational mobility outcomes movers tend to experience. One explanation may be that movers simply migrate to departments where wages are higher. To investigate this channel, we compute two statistics: (i) the mean family household wage rank in the origin department, and (ii) the mean child household income rank in the destination department. Figure 9 displays the average of these two statistics for movers for each ventile of the parent household wage rank distribution. There are three takeaways from this figure. First, the difference in average income rank in the destination and origin departments is highest at the top and bottom of the parent income distribution. Second, these differences are relatively small, reaching at most 3 percentile ranks for the top ventile. Third, the origin and destination departments of movers from the middle of the parent income distribution have very similar average income ranks. Put in parallel with the slight monotonic decrease in the gains from geographic mobility along the parent income rank distribution, it seems that these gains are not only due to individuals moving to higher-income departments.

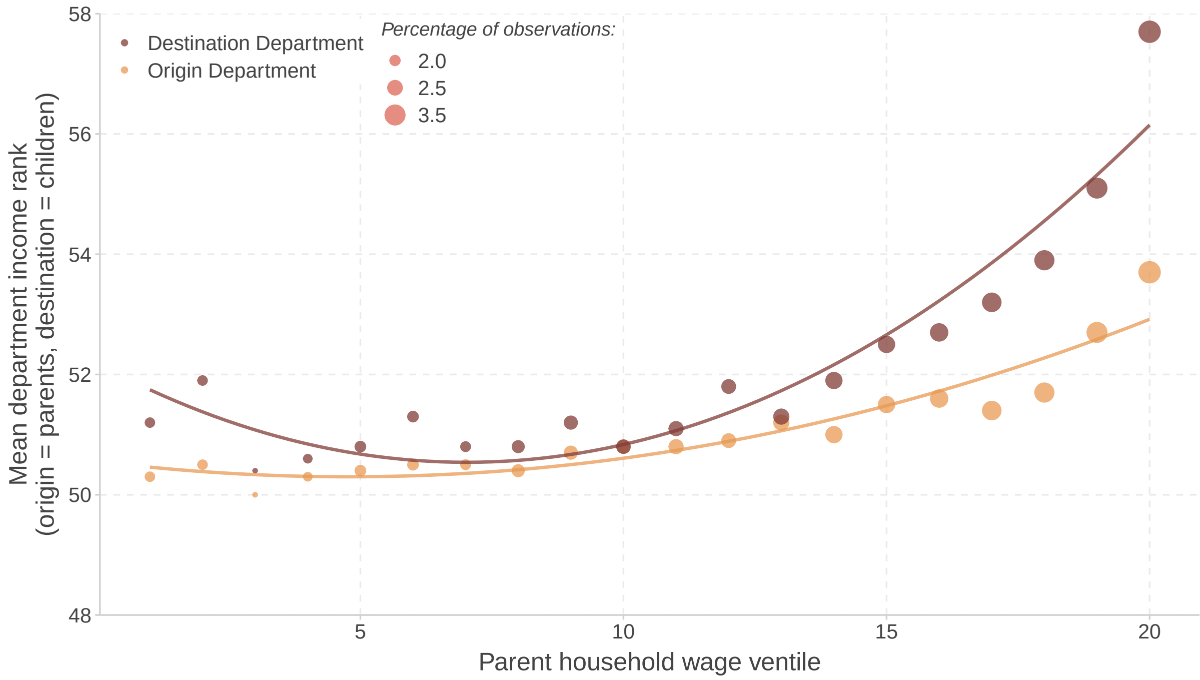



**Figure 9:** Mean Income Rank of Origin and Destination Departments of Movers

*Notes:* This figure represents the conditional expectation of income rank with respect to parent household wage rank for movers, separately by origin and destination departments. Origin department mean income rank is computed as the average income rank of residents in the parent sample, while destination mean income rank is computed as the average income rank of residents in the child sample. See Figures 1 and 8's notes for details on data, sample and income definitions.

Another way to test this hypothesis consists in comparing the conditional expectation functions of movers and stayers ranked both at the *national* and *department* level. Indeed, ranking individuals at the national level allows individuals born to parents who earn the median income of their department to be upward mobile by earning the median income of a higher-income department in adulthood. This channel can be removed by ranking individuals and their parents within departments. When doing so, movers can only be more intergenerationally mobile than stayers if they reach income ranks in their adulthood department that are further away from the rank of their parents in their childhood department. Finding no expected gains associated with geographic mobility when ranking individuals according to their department income distribution would suggest that the expected increase in income rank associated with mobility is fully driven by movers ending up in higher-income departments, but reaching on expectation a local income rank in their destination department that is not further away from that of their parents, relative to stayers.

The regression results of equation (6) using percentile ranks computed at the department level instead of at the national level are reported in Appendix Table F.12.[53] When considering ranks in the department distribution, the gap between the conditional expectation functions of movers and stayers shrinks but does not vanish completely. While the expected national-rank increase associated with mobility amounts to 5.95, it drops to 3.35 when considering local ranks. This suggests that the intergenerational mobility gains associated with geographic mobility are partly attributable to movers locating in higher-income departments in adulthood relative to stayers, but also to movers reaching local ranks in their adulthood department that are further away from the rank of their parents in the childhood department.

**The Role of Mobility Toward Richer Departments at the Individual Level.** Geographic mobility patterns between low- and high-income departments only partially explain the gap between movers and stayers at the aggregate level. Yet, differences in origin and destination department characteristics may be decisive at the individual level. In particular, the few movers transitioning to a high-income department may greatly benefit from this geographic mobility. To investigate this hypothesis we classify destination departments in three groups according to the average income rank of their residents in the child cohort: (i) *low-income*, destination departments with an average income rank below 50 (71 departments - 74% of movers), (ii) *medium-income*, those with an average income rank between 50 and 65 (20 departments and overseas departments - 21% of movers), and (iii) *high-income*, those with an average income rank above 65 (4 departments and foreign countries - 5% of movers).

Figure 10 shows the conditional expectation of child income rank with respect to parent income ventile for the three destination department categories and for stayers. Results of the corresponding regression are reported in Appendix Table F.13. Except for the top ventiles, the CEFs of movers by destination department category are virtually parallel. Movers thus experience the same rank-rank correlation regardless of the income category of their destination department. It is slightly lower than that of stayers. However, movers do not share the same absolute upward mobility, which increases with the average income of the destination department. These parallel shifts are such that the expected income rank of a mover from the bottom of the parent income distribution to a high-income department is around the same as the expected in-

[53] Appendix Figure E.15 shows the corresponding conditional expectation functions.

come rank of a stayer from the 75$^{th}$ percentile of the parental income distribution. Still, such transitions are the exception: most movers to high-income departments come from high-income families, while low-income movers go predominantly to low- or medium-income departments. This is coherent with the fact that the gap observed at the aggregate level between stayers and movers in Figure 8 is moderate.

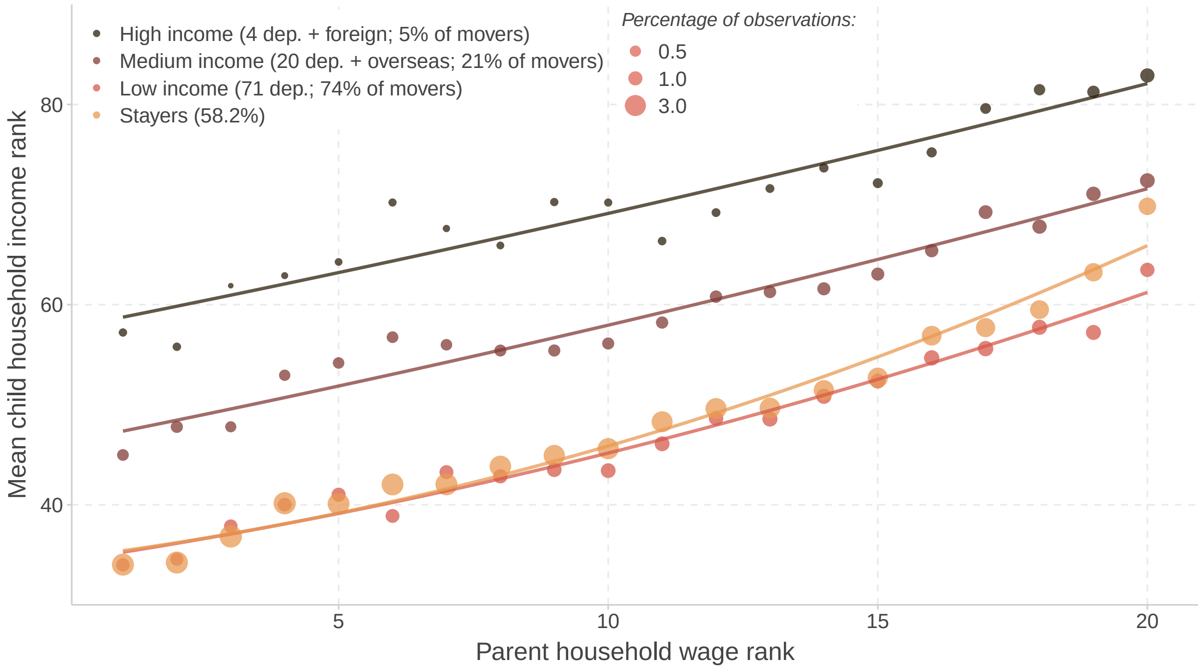



**Figure 10:** Mean Child Income Rank by Destination Department Mean Income

*Notes:* This figure represents the conditional expectation of child household income rank with respect to parent household wage rank for stayers and for movers to departments of different mean income categories. Low income destination departments are destination departments with an average income rank below 50, medium income are those with an average income rank between 50 and 65, and high income are those with an average income rank above 65. See Figures 1 and 8's notes for details on data, sample and income definitions.

Another noteworthy finding is that expected income ranks are essentially the same for movers to low-income departments as for stayers, highlighting the potential role of the destination department's characteristics in generating upward intergenerational mobility for movers. All these findings combine self-selection and causal effects, and we leave the disentangling of these two channels for future research.

# 7 Conclusion

France is an interesting case study for intergenerational income mobility considering its relatively modest income inequality and the specificity of its higher education system. Yet, it has been the focus of few studies due to important data limitations. We use administrative data to provide an overview of intergenerational income mobility in France for individuals born in 1972-1981. Relative to existing studies, the richness of these data enables us to apply two-sample two-stage least squares using a much larger set of individual characteristics, and to extensively assess the robustness of the resulting estimates.

Moreover, we provide the first estimates of the rank-rank correlation and transition matrix for France, and conduct a comparative analysis with other countries for which such statistics are available. Our results reveal that France exhibits a relatively strong intergenerational income persistence at the national level. It ranks among the highest in OECD countries, with Italy and the United States, and far from Switzerland, Australia, and the Scandinavian countries.

This high intergenerational income persistence at the national level hides substantial geographic heterogeneity across departments. We observe about as much variation across French departments as we do across across countries. Intergenerational persistence appears to be particularly high in the North and South, and relatively low in the Western part of the country. Yet, only *absolute* mobility, as opposed to *relative* mobility, significantly correlates with local characteristics.

We also provide novel descriptive evidence on a new mechanism that could explain some features of intergenerational mobility: geographic mobility. We find that the difference in expected income ranks between geographically mobile individuals and stayers is large and slightly decreasing in parent income. This difference appears not to be solely due to individuals moving to higher income departments but to be also the result of individuals moving up the local income rank ladder. Destination departments are on average characterized by higher income levels than origin departments only at the tails of the parent income distribution. However, regardless of parent income rank, conditional on moving the absolute upward mobility gains associated with moving to a higher-income department appear to be large and increasing with average income in the destination department. Even though not causal, we believe that these descriptive findings constitute promising avenues for future research to better understand intergenerational income mobility.

# References


**Abbas, Hicham, and Michaël Sicsic.** 2022. "Qui Gravit l'échelle Des Revenus Par Rapport à Ses Parents ?" INSEE Documents de Travail N° 2022 - F04, Paris.

**Acciari, Paolo, Alberto Polo, and Giovanni L Violante.** forthcoming. ""And Yet, It Moves": Intergenerational Mobility in Italy." *American Economic Journal: Applied Economics*.

**Alesina, Alberto, Stefanie Stantcheva, and Edoardo Teso.** 2018. "Intergenerational Mobility and Preferences for Redistribution." *American Economic Review*, 108(2): 521–554.

**Barbieri, Teresa, Francesco Bloise, and Michele Raitano.** 2020. "Intergenerational Earnings Inequality: New Evidence From Italy." *Review of Income and Wealth*, 66(2): 418–443.

**Bellégo, Christophe, David Benatia, and Louis-Daniel Pape.** 2021. "Dealing with Logs and Zeros in Regression Models." SSRN Electronic Journal.

**Björklund, Anders, and Markus Jäntti.** 1997. "Intergenerational Income Mobility in Sweden Compared to the United States." *American Economic Review*, 87(5): 1009–1018.

**Björklund, Anders, Jesper Roine, and Daniel Waldenström.** 2012. "Intergenerational Top Income Mobility in Sweden: Capitalist Dynasties in the Land of Equal Opportunity?" *Journal of Public Economics*, 96(5): 474–484.

**Black, Sandra E., and Paul J. Devereux.** 2011. "Recent Developments in Intergenerational Mobility." In *Handbook of Labor Economics*. Vol. 4, 1487–1541. Elsevier.

**Bonneau, Cécile, and Sébastien Grobon.** 2022. "Unequal Access to Higher Education Based on Parental Income: Evidence from France." World Inequality Lab 2022/01, Paris.

**Bratberg, Espen, Jonathan Davis, Bhashkar Mazumder, Martin Nybom, Daniel D. Schnitzlein, and Kjell Vaage.** 2017. "A Comparison of Intergenerational Mobility Curves in Germany, Norway, Sweden, and the US." *The Scandinavian Journal of Economics*, 119(1): 72–101.

**Carmichael, Fiona, Christian K. Darko, Marco G. Ercolani, Ceren Ozgen, and W. Stanley Siebert.** 2020. "Evidence on Intergenerational Income Transmission Using Complete Dutch Population Data." *Economics Letters*, 189: 108996.

**Chetty, Raj, and Nathaniel Hendren.** 2018. "The Impacts of Neighborhoods on Intergenerational Mobility II: County-Level Estimates." *The Quarterly Journal of Economics*, 133(3): 1163–1228.

**Chetty, Raj, John N Friedman, Emmanuel Saez, Nicholas Turner, and Danny Yagan.** 2020. "Income Segregation and Intergenerational Mobility Across Colleges in the United States." *The Quarterly Journal of Economics*, 135(3): 1567–1633.

**Chetty, Raj, Nathaniel Hendren, Patrick Kline, and Emmanuel Saez.** 2014. "Where Is the Land of Opportunity? The Geography of Intergenerational Mobility in the United States." *The Quarterly Journal of Economics*, 129(4): 1553–1623.

**Chuard-Keller, Patrick, and Veronica Grassi.** 2021. "Switzer-Land of Opportunity: Intergenerational Income Mobility in the Land of Vocational Education." SSRN Electronic Journal.

**Corak, Miles.** 2013. "Income Inequality, Equality of Opportunity, and Intergenerational Mobility." *Journal of Economic Perspectives*, 27(3): 79–102.

**Corak, Miles.** 2016. "Inequality from Generation to Generation: The United States in Comparison." IZA Discussion Paper 9929, Bonn, Germany.

**Corak, Miles.** 2020. "The Canadian Geography of Intergenerational Income Mobility." *The Economic Journal*, 130(631): 2134–2174.

**Couch, Kenneth A., and Dean R. Lillard.** 1998. "Sample Selection Rules and the Intergenerational Correlation of Earnings." *Labour Economics*, 5(3): 313–329.

**Dahl, Molly, and Thomas DeLeire.** 2008. "The Association between Children's Earnings and Fathers' Lifetime Earnings: Estimates Using Administrative Data." Institute for Research on Poverty No. 1342-08.

**Deutscher, Nathan, and Bhashkar Mazumder.** 2020. "Intergenerational Mobility across Australia and the Stability of Regional Estimates." *Labour Economics*, 66: 101861.

**Eriksen, Jesper.** 2018. "Finding the Land of Opportunity: Intergenerational Mobility in Denmark." Aalborg University.

**Grawe, Nathan D.** 2006. "Lifecycle Bias in Estimates of Intergenerational Earnings Persistence." *Labour Economics*, 13(5): 551–570.

**Haider, Steven, and Gary Solon.** 2006. "Life-Cycle Variation in the Association between Current and Lifetime Earnings." *American Economic Review*, 96(4): 1308–1320.

**Heidrich, Stefanie.** 2017. "Intergenerational Mobility in Sweden: A Regional Perspective." *Journal of Population Economics*, 30(4): 1241–1280.

**Helsø, Anne-Line.** 2021. "Intergenerational Income Mobility in Denmark and the United States." *The Scandinavian Journal of Economics*, 123(2): 508–531.

**Huang, Xiao, Shoujun Huang, and Ailun Shui.** 2021. "Government Spending and Intergenerational Income Mobility: Evidence from China." *Journal of Economic Behavior & Organization*, 191: 387–414.

**Inoue, Atsushi, and Gary Solon.** 2010. "Two-Sample Instrumental Variables Estimators." *The Review of Economics and Statistics*, 92(3): 557–561.

**Jerrim, John, Álvaro Choi, and Rosa Simancas.** 2016. "Two-Sample Two-Stage Least Squares (TSTSLS) Estimates of Earnings Mobility: How Consistent Are They?" *Survey Research Methods*, 10(2): 85–102.

**Kalambaden, Preetha, and Isabel Z Martınez.** 2021. "Intergenerational Mobility Along Multiple Dimensions Evidence from Switzerland."

**Laliberté, Jean-William.** 2021. "Long-Term Contextual Effects in Education: Schools and Neighborhoods." *American Economic Journal: Economic Policy*, 13(2): 336–377.

**Landersø, Rasmus, and James J. Heckman.** 2017. "The Scandinavian Fantasy: The Sources of Intergenerational Mobility in Denmark and the US." *The Scandinavian Journal of Economics*, 119(1): 178–230.

**Lefranc, Arnaud.** 2018. "Intergenerational Earnings Persistence and Economic Inequality in the Long Run: Evidence from French Cohorts, 1931-75." *Economica*, 85(340): 808–845.

**Lefranc, Arnaud, and Alain Trannoy.** 2005. "Intergenerational Earnings Mobility in France: Is France More Mobile than the US?" *Annales d'Économie et de Statistique*, , (78): 57–77.

**Lefranc, Arnaud, Nicolas Pistolesi, and Alain Trannoy.** 2009. "Equality of Opportunity and Luck: Definitions and Testable Conditions, with an Application to Income in France." *Journal of Public Economics*, 93(11-12): 1189–1207.

**Mayer, Susan E., and Leonard M. Lopoo.** 2008. "Government Spending and Intergenerational Mobility." *Journal of Public Economics*, 92(1-2): 139–158.

**Mazumder, Bhashkar.** 2005. "Fortunate Sons: New Estimates of Intergenerational Mobility in the United States Using Social Security Earnings Data." *Review of Economics and Statistics*, 87(2): 235–255.

**Mazumder, Bhashkar.** 2016. "Estimating the Intergenerational Elasticity and Rank Association in the United States: Overcoming the Current Limitations of Tax Data." In *Research in Labor Economics*. Vol. 43, , ed. Lorenzo Cappellari, Solomon W. Polachek and Konstantinos Tatsiramos, 83–129. Emerald Group Publishing Limited.

**Mazumder, Bhashkar, and Nathan Deutscher.** forthcoming. "Measuring Intergenerational Income Mobility: A Synthesis of Approaches." *Journal of Economic Literature*.

**Nybom, Martin, and Jan Stuhler.** 2016. "Heterogeneous Income Profiles and Lifecycle Bias in Intergenerational Mobility Estimation." *Journal of Human Resources*, 51(1): 239–268.

**Nybom, Martin, and Jan Stuhler.** 2017. "Biases in Standard Measures of Intergenerational Income Dependence." *Journal of Human Resources*, 52(3): 800–825.

**OECD.** 2018. *A Broken Social Elevator? How to Promote Social Mobility.* OECD.

**O'Neill, Donal, Olive Sweetman, and Dirk Van de gaer.** 2007. "The Effects of Measurement Error and Omitted Variables When Using Transition Matrices to Measure Intergenerational Mobility." *Journal of Economic Inequality*, 5(2): 159–178.

**Robert-Bobée, Isabelle, and Natacha Gualbert.** 2021. "L'échantillon Démographique Permanent : En 50 Ans, l'EDP a Bien Grandi !" In *Courrier Des Statistiques N6*. Paris:INSEE.

**Solon, Gary.** 1992. "Intergenerational Income Mobility in the United States." *American Economic Review*, 82(3): 393–408.

**Soria Espín, Javier.** 2022. "Intergenerational Mobility, Gender Differences and the Role of Out-Migration: New Evidence from Spain." 98.

**Zimmerman, David J.** 1992. "Regression Toward Mediocrity in Economic Stature." *American Economic Review*, 82(3): 409–429.

# A Data - Details

The Permanent Demographic Sample (EDP) is a panel of individuals which the French statistical office, INSEE, started in 1968. It combines several administrative data sources on individuals born on the first four days of October.[54] Individuals born on one of these days are called EDP individuals. The EDP gathers data from 5 administrative sources: (i) civil registers since 1968; (ii) population censuses since 1968 (exhaustive in 1968, 1975, 1982, 1990 and 1999, and yearly rotating 20% random samples since 2004); (iii) the electoral register since 1990; (iv) the *All Employee Panel* since 1967; and (v) tax returns since fiscal year 2011.

Each time an individual born on the first four days of October appears in one of these administrative datasets, the information contained in it is added to their individual identifier in the EDP. Therefore all these datasets can be matched together using a common individual identifier. For our analysis we use data from civil registers, the 1990 census, the All Employee Panel and tax returns. We describe each data source in detail below.

**Civil Registers.** They contain information from birth certificates of EDP individuals and their children, as well as death and marriage certificates of EDP individuals, since 1968. We use birth certificates of EDP individuals and their children which include the child's gender, date and place of birth, and information on each parent including date and place of birth, nationality and occupation. There are no data breaks or missing certificates for the years under study (1972-1981).

**1990 Census.** It contains socio-demographic information about EDP individuals, as well as, though to a lesser extent, about members of their household. These include the individual's date and place of birth, nationality, education, occupation, marital status, household structure, dwelling characteristics, building when relevant, and municipality.

**All Employee Panel.** It combines two sources of data: the annual declarations of social data (*déclarations annuelles des données sociales* - DADS) and data on central government employees (*fichiers de paie des agents de l'état* - FPE). All businesses are obliged to annually communicate the declarations of social data about their employees to a network of private organizations (*Unions de recouvrement des cotisations de sécurité sociale et d'allocations familiales* - URSSAF) coordinated by a government agency (*Agence centrale des organismes de sécurité sociale* - ACOSS). The All Employee Panel data are reported at the worker-year level, aggregated by INSEE from data at the worker-firm-year level. As such, annual pretax wage and annual hours worked correspond to the sum over all the individual's salaried activities. The job characteristics correspond to the year's "main" job, that is the job for which the pay period was the longest and, in case of a tie, the job with the highest wage.

Between 1967 and 2001, data is only available for individuals born on an even year. The scope of workers covered by the All Employee Panel has varied over time. Since 1967 in metropolitan France, all private sector employees, except those in the agricultural sectors, and including employees of public enterprises, are covered. The hospital

[54]The EDP selection criterion has progressively widened to include individuals born on the first days of January, April, and July.

public service is integrated in 1984, the state civil service and local authorities in 1988.[55] The agricultural sector and overseas territories are included in 2002, and employees of private employers in 2009. Unemployment insurance is included from 2008 onwards. Lastly, because of increased workload due to the population censuses of 1982 and 1990, the All Employee Panel data were not compiled by INSEE in 1981, 1983 and 1990.

**Tax Returns.** They are compiled using housing and income tax forms filed for incomes earned from 2010 to 2016. In particular, household-level tax returns information is constructed based on dwellings where an EDP individual is known either from the income tax return or from the principal housing tax (*taxe d'habitation principale*). The location of the individual is that declared on January 1$^{st}$ of the fiscal declaration year. Income variables are available at the household-level as well as at the individual level. Since the information is gathered based on living in the same dwelling, household income is computed not only for couples who file their taxes jointly, but also for couples who live together, an increasingly common arrangement. This departs from existing studies based on tax returns data which can only assign households based on marital status (Chetty et al., 2014). The scope of fiscal households excludes individuals living in collective structures (retirements homes, religious communities, student accommodations, prisons, etc.) as well as those most in distress, who live in precarious housing (worker hostels, etc.) or are homeless.

[55]France Télécom and La Poste employees appear only in 1988 as well. See Appendix B.1 for a robustness check to this public sector coverage evolution.

# B Additional Robustness

This Appendix provides additional robustness checks to those presented in the body of the paper.

## B.1 Sensitivity to Data Coverage

We ensure our results are not affected by the fact that civil servants are only observed from 1988 onwards by estimating the first-stage regression computing synthetic parents' on post-1988 wages only, still restricting to when they are between 35 ad 45 years old. Appendix Figure B.1 displays the results from this check. The results are largely unaffected.

## B.2 Alternative First-Stage Estimation

The parent income predictions we use to palliate French data limitations are central to our analysis. It is of primary importance that the first stage of the two-step strategy we rely on is reliable. We make sure that this first stage does not spuriously drive

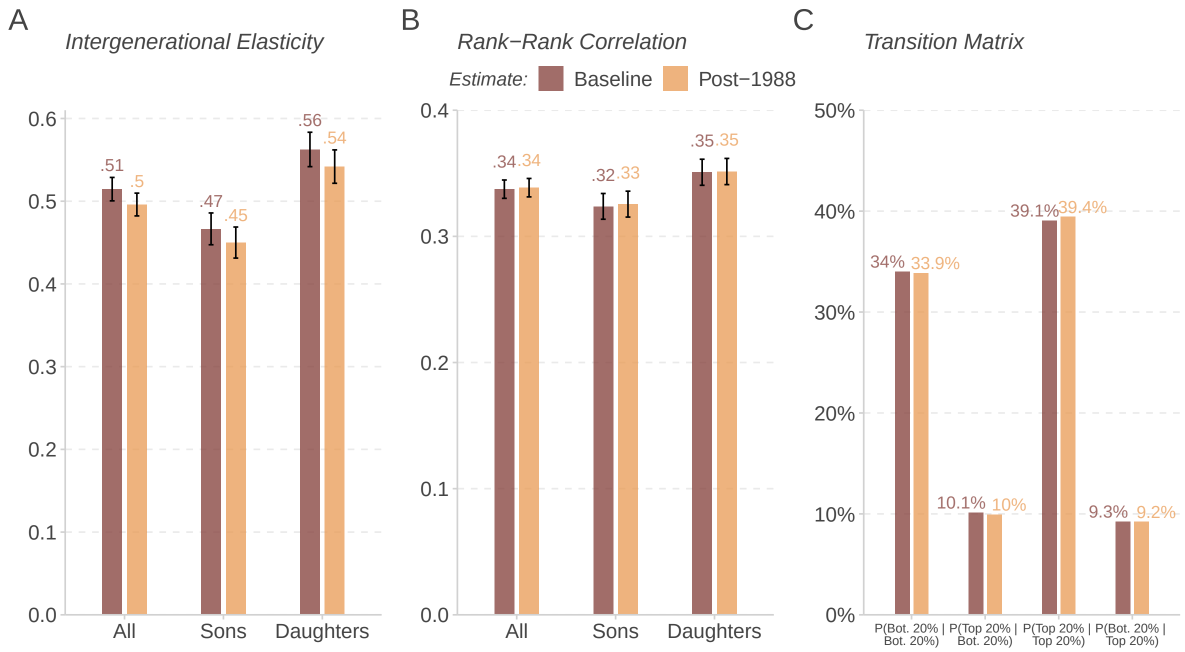



**Figure B.1:** Robustness of Baseline Estimates to Computing Synthetic Parent Incomes only on Post-1988 Data

*Notes:* This figure assesses the robustness of our baseline intergenerational income mobility estimates to computing synthetic parents' incomes only on post-1988 data. The All Employee Panel from which synthetic parents' wages are observed did not cover civil servants prior to 1988 (see Appendix Section A for details). The graph presents the baseline estimates (Baseline) to those obtained when synthetic parent incomes are defined as average wage between 35-45 using only post-1988 wages (Post-1988). All results pertain to parent and child incomes being defined at the household level. The results for the transition matrix correspond to the sample pooling sons and daughters. See Section 3 for details on data, sample and income definitions.

the results in one way or another by evaluating its sensitivity to relaxing parametric assumptions and varying both the set of instruments and sample restrictions.

We make use of semi- and non-parametric models to elicit potential misspecifications in the first stage. The baseline specification of the first stage is of the form $y = \beta X + \varepsilon$, where $y$ is the log of parent lifetime income and $X$ is a set of $k$ predictors. OLS would not account for interactions between predictors nor for non-linearities in the relationship between $X$ and $y$ unless they are explicitly modeled. Fully non-parametric methods of the form $y = m(X) + \varepsilon$ would capture both interactions and non-linearities that may help reduce the out-of-sample MSE. Obtaining a lower MSE and significantly different second-stage estimates with non-parametric models than with OLS would suggest that non-modeled non-linearities, interactions, or both, influence the resulting intergenerational mobility estimates.

We implement this test using three machine learning methods: (i) a generalized additive model (GAM) of the form $y = m_1(x_1) + m_2(x_2) + ... + m_k(x_k) + \varepsilon$ which accounts for non-linearities but not for interactions unless explicitly specified, (ii) a gradient boosted regression tree, that is a high-dimensional combination of sequentially grown regression trees, and (iii) the ensemble method, which consists in taking the average of the predictions from each model weighted in a way that minimizes the out-of-sample MSE.

Appendix Figure B.2 compares the intergenerational mobility estimates and out-of-sample MSE resulting from these three methods using our baseline child and parent income definitions. We do not observe significant differences in MSE between the different prediction methods. The resulting mobility estimates are virtually the same for OLS, GAM and the ensemble method, and slightly smaller for boosted trees. This suggests that conditional on the set of predictors we use, using more flexible estimation methods does not lead to better income predictions and different estimates than using an additive OLS specification.

The other dimension to consider is the set of variables included in the first stage, notably because it has been shown that inadequate instruments could yield inconsistent estimates (Jerrim, Choi and Simancas, 2016). Appendix Figure B.3 documents the sensitivity of IGE and RRC estimates to the set of predictors used in the first-stage estimation. We do not assess the sensitivity of the transition matrices because for those measures, the accuracy of the prediction matters more and therefore simple prediction models will necessarily be inadequate. We estimate the IGE and RRC for adding each of the following predictors sequentially (all measured in 1990): education (8 cat.), 2-digit occupation (42 categories), a group of demographic characteristics (age, French nationality dummy, country of birth (6 cat.), and household structure (6 cat.)) and a group of municipality-level characteristics (unemployment rate, share of single mothers, share of foreigners, population, and population density). Since relying on a single variable with less than 100 categories induces some income values to span over several percentiles, parents with a given predicted income are attributed the average rank of individuals earning that level of income. Lastly, we also report the $R^2$ and root mean squared error (RMSE), computed as the average from 5-fold cross-validation.

We find that the IGE is 0.66 when using only education as the first-stage predictor, consistent with a point already made in the literature that using only education as a predictor is likely to yield inflated estimates of the IGE. Once 2-digit occupation is included in the first-stage, adding other demographic or city-level characteristics has no effect on the estimates. Indeed, as can be seen from the adjusted $R^2$, most of the pre-

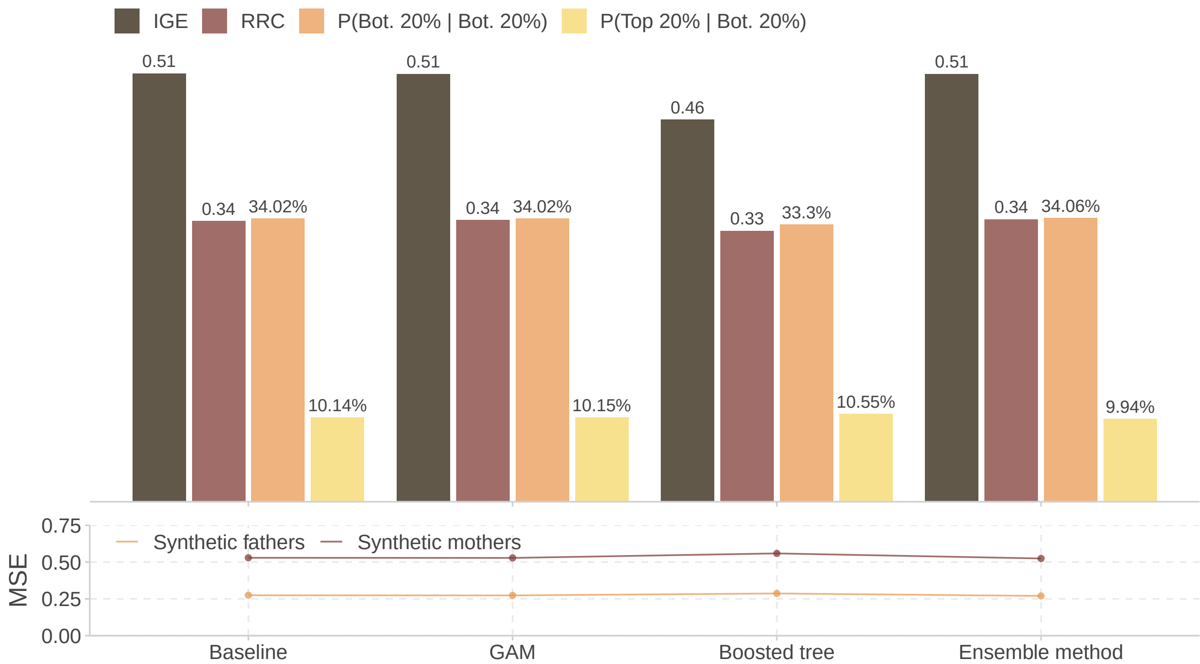



**Figure B.2:** Robustness to Machine Learning Prediction

*Notes:* This figure assesses the robustness of our baseline intergenerational income mobility estimates to increasingly flexible first-stage prediction models. Each bar represents the magnitude of the estimate of the corresponding color estimated using the first-stage model indicated on the x-axis. The first set of estimates are the baseline estimates obtained using OLS. The three other sets are obtained using increasingly flexible models: generalized additive models (GAM), gradient boosted regression trees, and the ensemble method. The connected dots represent the average out-of-sample MSEs of the associated prediction models, estimated using 5-fold cross-validation. See Figure 1's notes for details on data, sample and income definitions.

dictive power actually comes from the 2-digit occupation variable. The RRC appears remarkably unchanged by the set of first-stage predictors used, at 0.32 with only education and 0.33 with all variables. This appears once more to be a strength of the RRC in the TSTSLS context.

## B.3 Lifecycle and Attenuation Bias

**Child Lifecycle Bias - Constant Sample of Children.** To overcome the issue related to changes in Figure 4's underlying sample of children, we reproduce the individual wage estimates using the All Employee Panel keeping the sample of children constant. To do so we restrict to children born in 1972 and 1974[56] for whom wages are observed every year between 25 and 43 years old and 25 and 41 years old respectively. Appendix Figure B.4 displays the results. Since the sample is kept constant throughout, the coefficients can be compared to one another and the change in magnitude can only be driven by the age at which child income is measured rather than sample composition. As in Figure 4, we find that measuring child income prior to age 30 seriously underestimating the IGE (panel A) and RRC (panel B), and overestimating (underestimating) bottom or top mobility (persistence) (panel C).

**Child and Parent Lifecycle Bias Jointly.** Child and parent lifecycle bias are typically assessed independently, as we do in the main body of the article. Yet they influence one another and it is instructive to estimate our measures of intergenerational persistence for each possible combination of synthetic parent and child age. Appendix Figure B.5 shows such estimates when child income is measured between ages 30 and 44, and synthetic parent income between ages 28 and 60.

**Child Attenuation "Bias".** Appendix Figure B.6 plots estimates of our persistence measures varying the number of child income observations from 1 to 7 between 35 and 45 years old, keeping the sample of children constant[57] (i.e. keeping only children with 7 household income observations). Due to this restriction only cohorts born between 1972 and 1975 are kept. In the same way as for parents, we control for lifecycle bias by centering the year in which child income is measured to 2013. In other words, one child income observation corresponds to income measured in 2013, two income observations corresponds to the average between 2012 and 2014, to average income between 2012 and 2014, etc. The results suggest that estimates are largely unaffected by increasing the number of child income observations.

**Parent Attenuation Bias - Constant Sample of Synthetic Parents.** We check whether the lack of change in intergenerational mobility measures with the number of synthetic parent income observations observed in Figure 6 could be due to the fact that the sample of synthetic parents varies throughout. We replicate those estimates restricting the sample of synthetic parents to those with all 11 income observations between 35 and 45 years old and estimating the intergenerational mobility measures by varying the num-

[56] We cannot include the 1973 cohort as the All Employee Panel income data are only available for individuals born an even year before 2001. This choice of cohorts is done to be able to measure their incomes after they are 40 years old.

[57] The sample varies ever so slightly for the IGE due to the number of negative or 0 incomes changing between years.

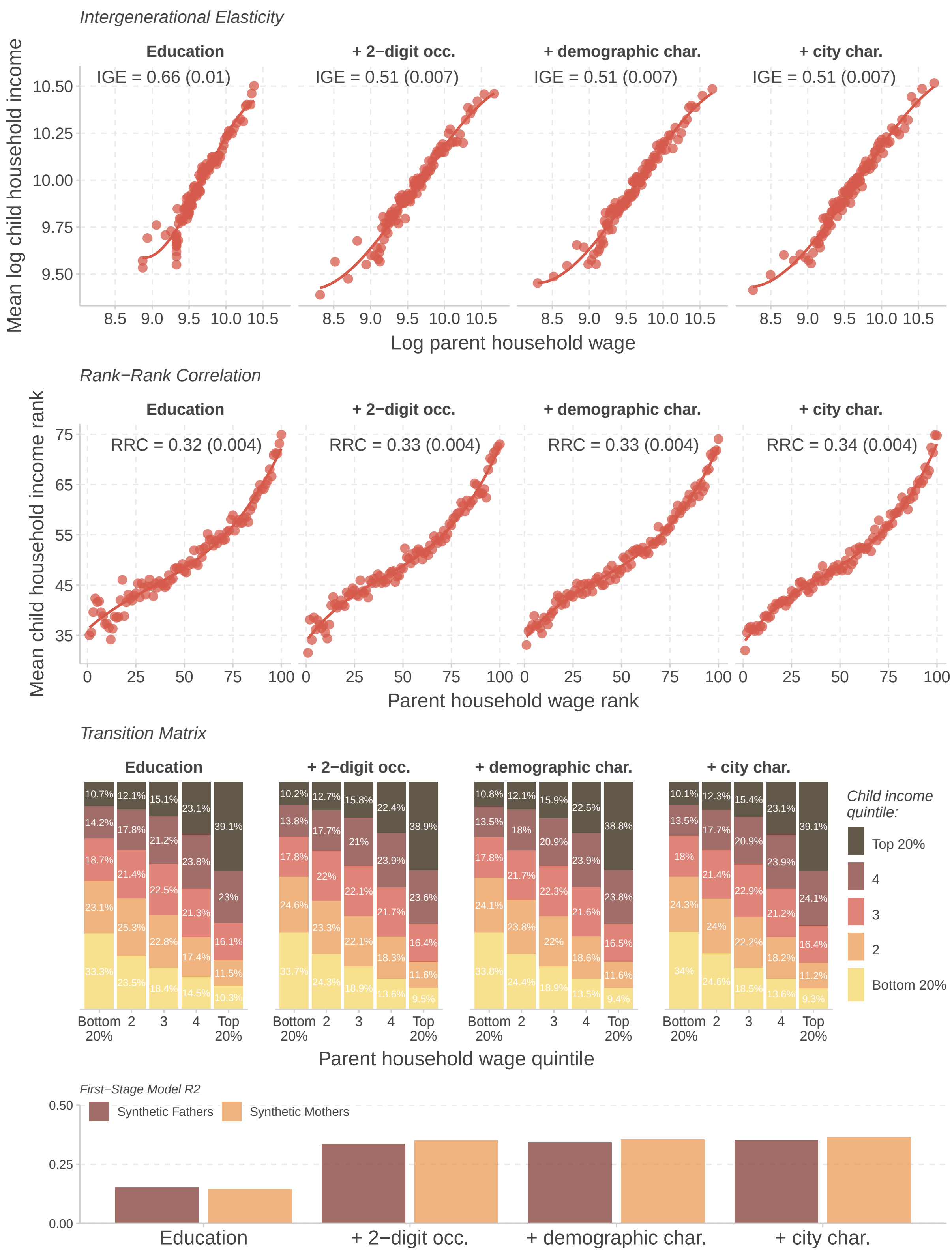



**Figure B.3:** Robustness of Baseline Estimates to Different First-Stage Predictors

*Notes:* This figure assesses the robustness of our baseline IGE, RRC and transition matrix estimates to variations in the set of first-stage predictors. Parent income is predicted separately for fathers and mothers using a set a of instruments that vary along the x-axis. We report the corresponding CEFs, along with the point estimates and the standard error in parenthesis. The bottom panel of the figure reports separately for synthetic fathers and mothers the $R^2$ associated with each first stage. See Figure 1's notes for details on data, sample and income definitions.

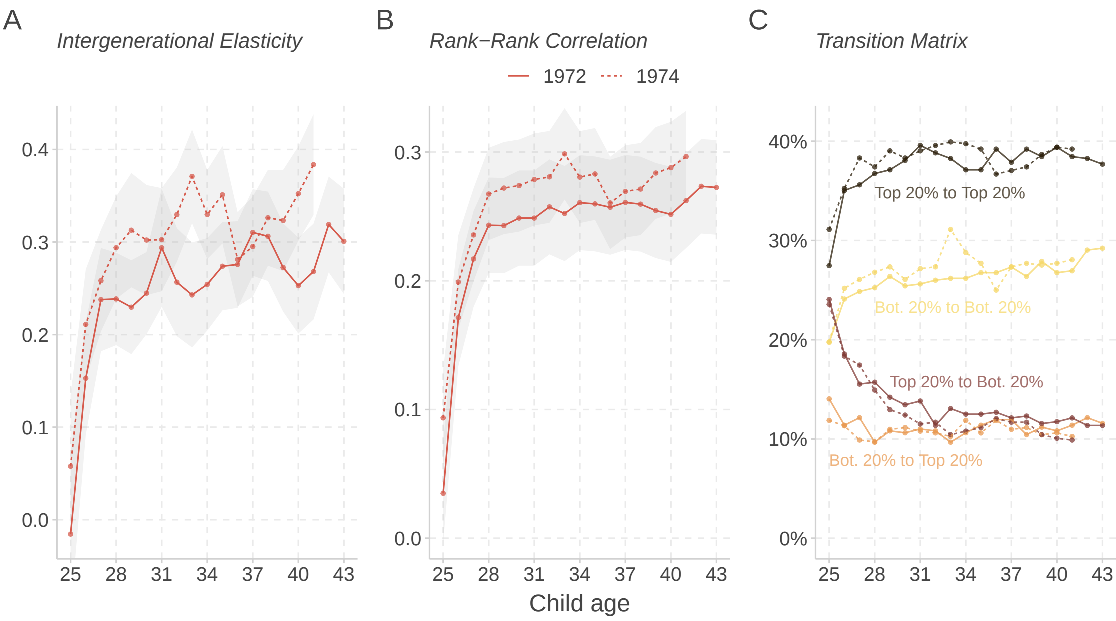
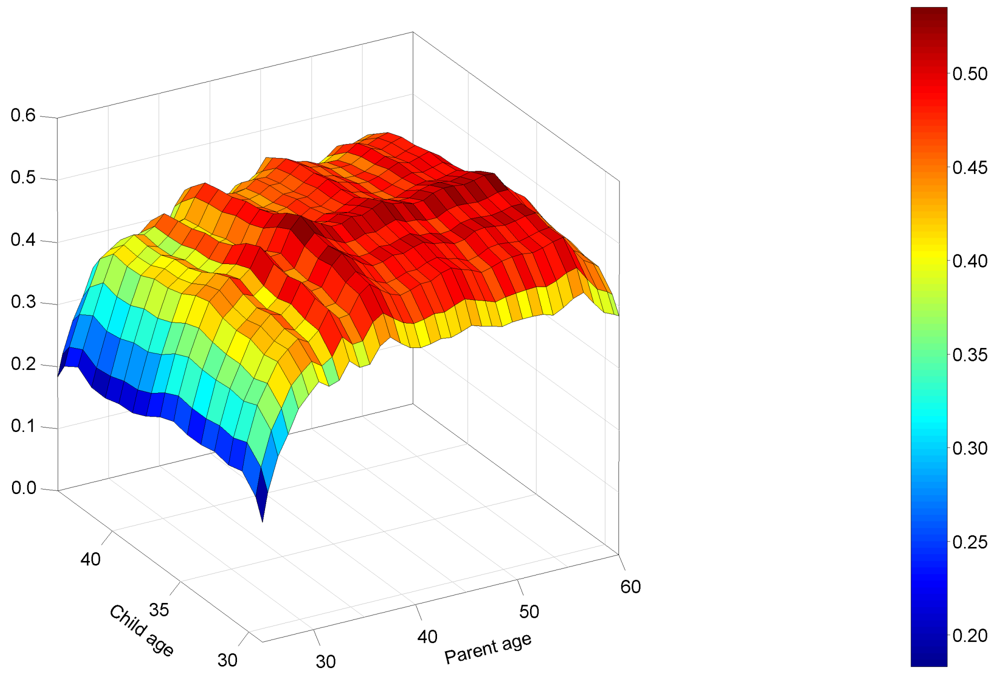
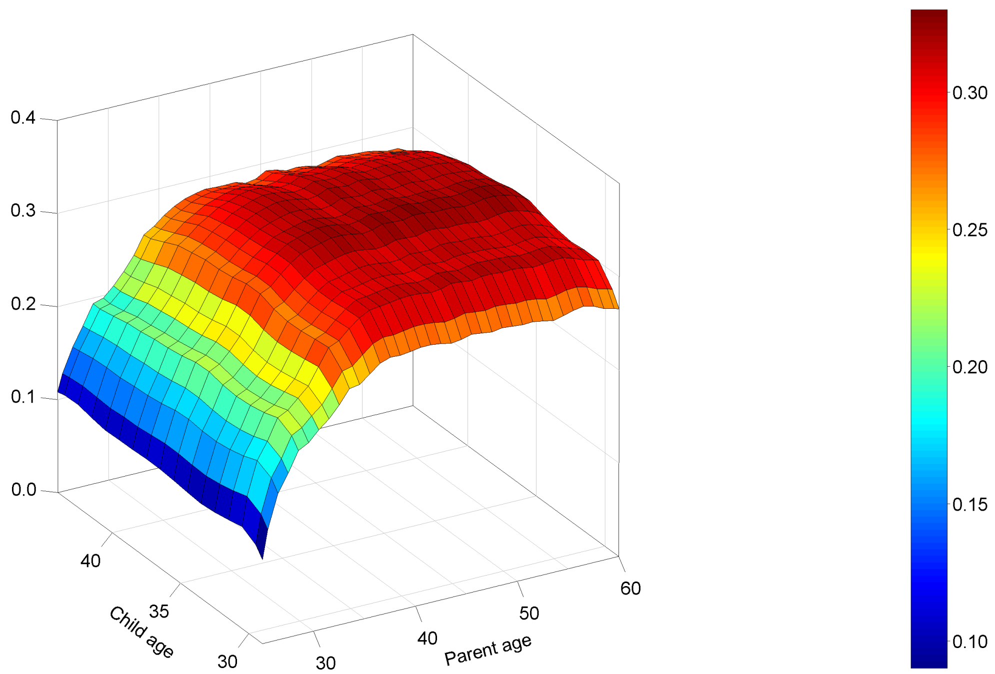



**Figure B.4:** Child Lifecycle Bias - 1972 and 1974 Cohorts (Constant Sample)

*Notes:* This figure assesses the robustness of our baseline intergenerational income mobility estimates presented in Figures 1 and 2 to changes in the age at which child income is measured, for children born in 1972 (solid line) and 1974 (dashed line). For both birth cohorts the sample is kept constant, that is only children with wages observed in the All Employee Panel at each age between 25 and 43 years old are retained. Shaded areas represent the 95% confidence interval. See Sections 3 and 4.4 for for details on data, sample and income definitions.




**(a)** Intergenerational Elasticity




**(b)** Rank-Rank Correlation

**Figure B.5:** Child and Parent Lifecycle Bias

*Notes:* This figure assesses the robustness of our baseline intergenerational income mobility estimates presented in Figure 1 to changes in the age at which child and synthetic parent incomes are measured. The sample of children and synthetic parents varies across ages. See Sections Figure 1's notes for for details on data, sample and income definitions.

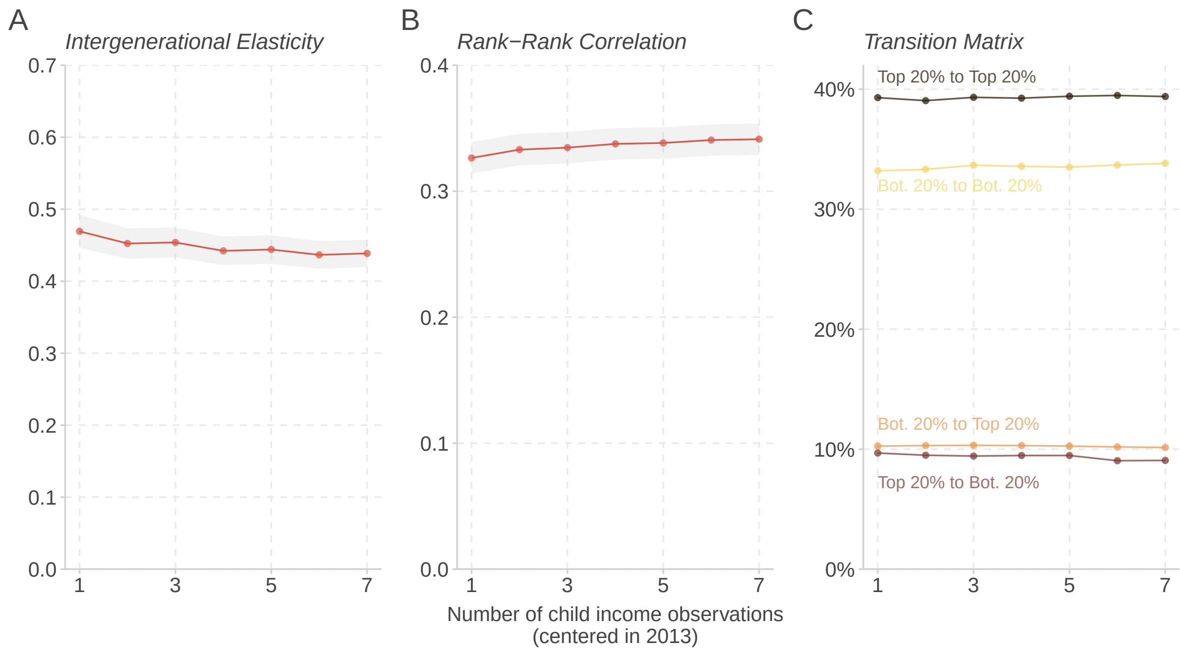



**Figure B.6:** Sensitivity to Number of Child Income Observations (Constant Sample)

*Notes:* This figure presents estimates of our persistence measures varying the number of child income observations from 1 to 7 between 35 and 45 years old, keeping the sample of children constant, i.e. keeping only children with 7 household income observations. (The sample varies ever so slightly for the IGE due to the number of negative or 0 incomes varying between years.) Due to this restriction only cohorts born between 1972 and 1975 are kept. We control for lifecycle bias by centering the year in which child income is measured to 2013. In other words, one child income observation corresponds to income measured in 2013, two income observations corresponds to the average of 2012 and 2014, three to average income between 2012 and 2014, etc. Shaded areas represent the 95% confidence interval. See Figure 1's notes for for details on data, sample and income definitions.

ber of income observations averaged in the first-stage regression (centered around 40 years old again). To do so, we impute wages in 1981, 1983 and 1990, for which the data are not available,[58] using the average wage between the previous and subsequent year only if both wages are observed. This enables us to have a consistent sample and increase the number of synthetic parents on which the predictions can be done.

Appendix Figure B.7 displays the results from this sensitivity analysis. The increase in the parent household wage IGE is much less marked, increasing from 0.618 when using one income observation to 0.693 when using all 11 observations (panel A). Our interpretation of this relatively modest increase is that averaging over at least 2 income observations as we do for our baseline estimate should suffice to not suffer from attenuation bias. Note that what matters in this figures is not how different the estimates are from our baseline estimate but rather the extent to which they vary with the number of synthetic parent income observations used. Indeed, the difference between our baseline IGE estimate and the estimates obtained are driven by the fact that the sample of synthetic parents for whom we observe all incomes between 35 and 45 years old is a highly non-representative sample, especially when it comes to mothers. In fact, we do not find any attenuation bias when restricting our analysis to fathers, suggesting all the variation in the IGE can be accounted for by changes in mothers' incomes predictions. As with the varying synthetic parent sample estimates, rank-based intergenerational mobility measures are significantly less sensitive to averaging over more income years, and the estimates found are very close to our baseline ones (panels B and C).

## B.4 Sensitivity to Income Distribution Tails.

Our baseline estimates may be sensitive to two main sample selection choices when it comes to the income distributions of parent and children: (i) how children with negative or zero incomes are treated; and (ii) how the top and bottom tails of both the parent and child income distributions are dealt with.

**Treatment of Zeros.** The first issue is particularly salient for the estimation of the intergenerational income elasticity due to the impossibility of taking the log of zero.[59] Many researchers simply discard such observations since they are likely not representative of lifetime income. Though this may potentially be the case if only short income time spans are available, we nonetheless evaluate how our baseline estimates of both the IGE and the RRC when replacing negative or zero child income values by 1 or 1,000 euros.

Appendix Figure B.8 shows estimates for the IGE and RRC when replacing income of children reporting negative or zero incomes by 1 euro or 1,000 euros, for different child income definitions. For our primary child income definition, household income, the estimates do not change due to there being very few children with negative or zero household income. However, for child income defined at the individual-level, for which the share of negative or zero incomes is more important, the IGE becomes highly sensitive to the recoding of such observations while the RRC remains unchanged. For

[58] As explained in Section 3, the 1982 and 1990 population censuses generated an extra workload which prevented INSEE from compiling the All Employee Panel data for these years.

[59] Various methods have been proposed to overcome this issue. Bellégo, Benatia and Pape (2021) describe such methods and propose a novel solution that can be applied to a variety of cases.

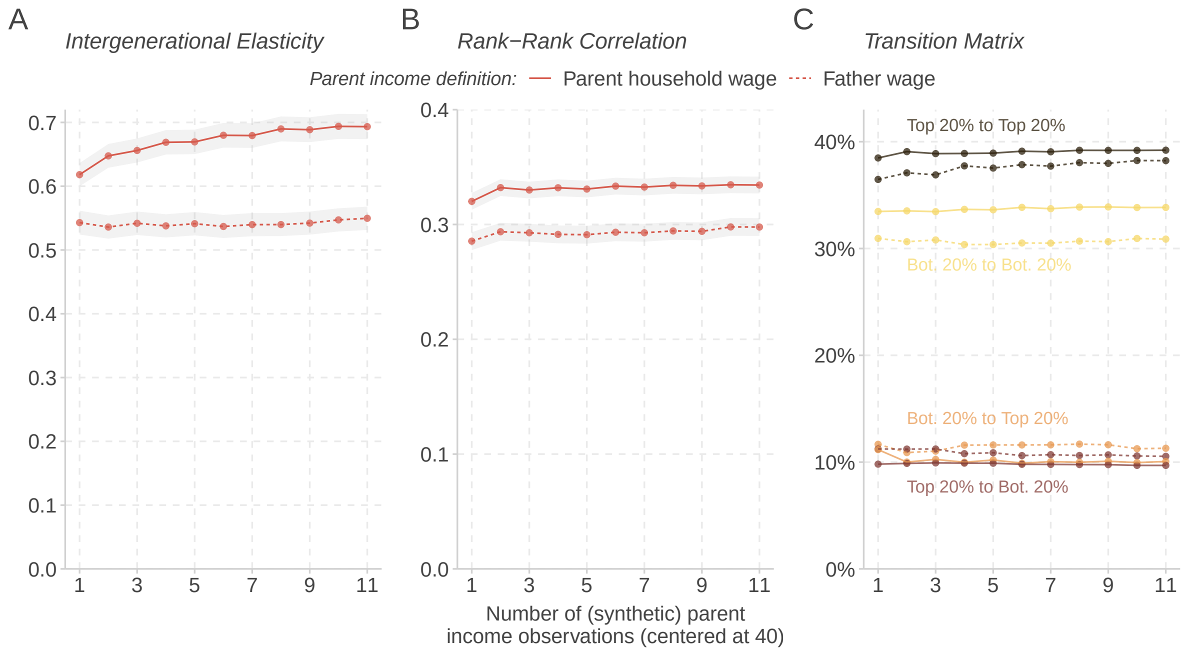



**Figure B.7:** Parent Attenuation Bias (Constant Sample of Synthetic Parents)

*Notes:* This figure assesses the robustness of our baseline intergenerational income mobility estimates to the number of income observations used to predict parent income, keeping the sample of synthetic parents constant. The sample of synthetic parents is thus restricted to those with all 11 income observations between 35 and 45 years old. While varying the number of parent income observations, we center the age range at 40 to control for lifecycle bias. Shaded areas represent the 95% confidence interval. See Figure 1's notes for details on data, sample and income definitions.

example, for individual child income, the IGE is 0.46 when zeros are dropped and 0.79 when they are recoded to 1 and 0.54 when recoded to 1,000. The RRC is entirely insensitive to such recoding as ranks are not altered by it.

**Top and Bottom Trimming.** The second issue relates to the treatment of top and bottom earners in both the parent and child income distributions. For the parent income distribution the choice can both be made in the prediction stage and in the second stage. Specifically, we assess how the IGE and RRC vary when trimming the top and/or bottom 1% to 5% and 10%. Appendix Figure B.9 displays the results of this sensitivity check. There are three main takeaways.

First, the IGE is significantly more sensitive to small changes in parent or child income distributions while the RRC remains relatively stable. For example, removing the top and bottom 1% of child incomes decreases the IGE from 0.515 to 0.411 while the RRC only decreases from 0.337 to 0.322. It does not seem desirable that a measure of intergenerational mobility should be so sensitive to excluding just 2% of children. Mathematically it can be linked to changes in the dispersion of the distribution of child incomes but conceptually it seems difficult to defend such responsiveness to minor sample changes.

Second, the IGE is quite strongly influenced by minor trimming in the first-stage prediction sample. For example, excluding the bottom and top 2% of synthetic parent incomes leads to an IGE of 0.584. Such exclusions are not uncommon in the literature though their relevance is unclear.[60] Meanwhile the RRC is once more remarkably untouched by first-stage parent income exclusions. In fact excluding the bottom and top 10% of synthetic parent incomes decreases the RRC to 0.333 (from 0.337). This appears to be an additional benefit of estimating the RRC when using with the TSTSLS method.

Third, for second-stage parent income trimming, the effects are relatively mild for both intergenerational mobility measures. This is very likely a consequence of the two-stage procedure which reduces the variance in parent incomes.

[60] For example, Barbieri, Bloise and Raitano (2020) exclude the top and bottom 1% of their sons and synthetic fathers' incomes.

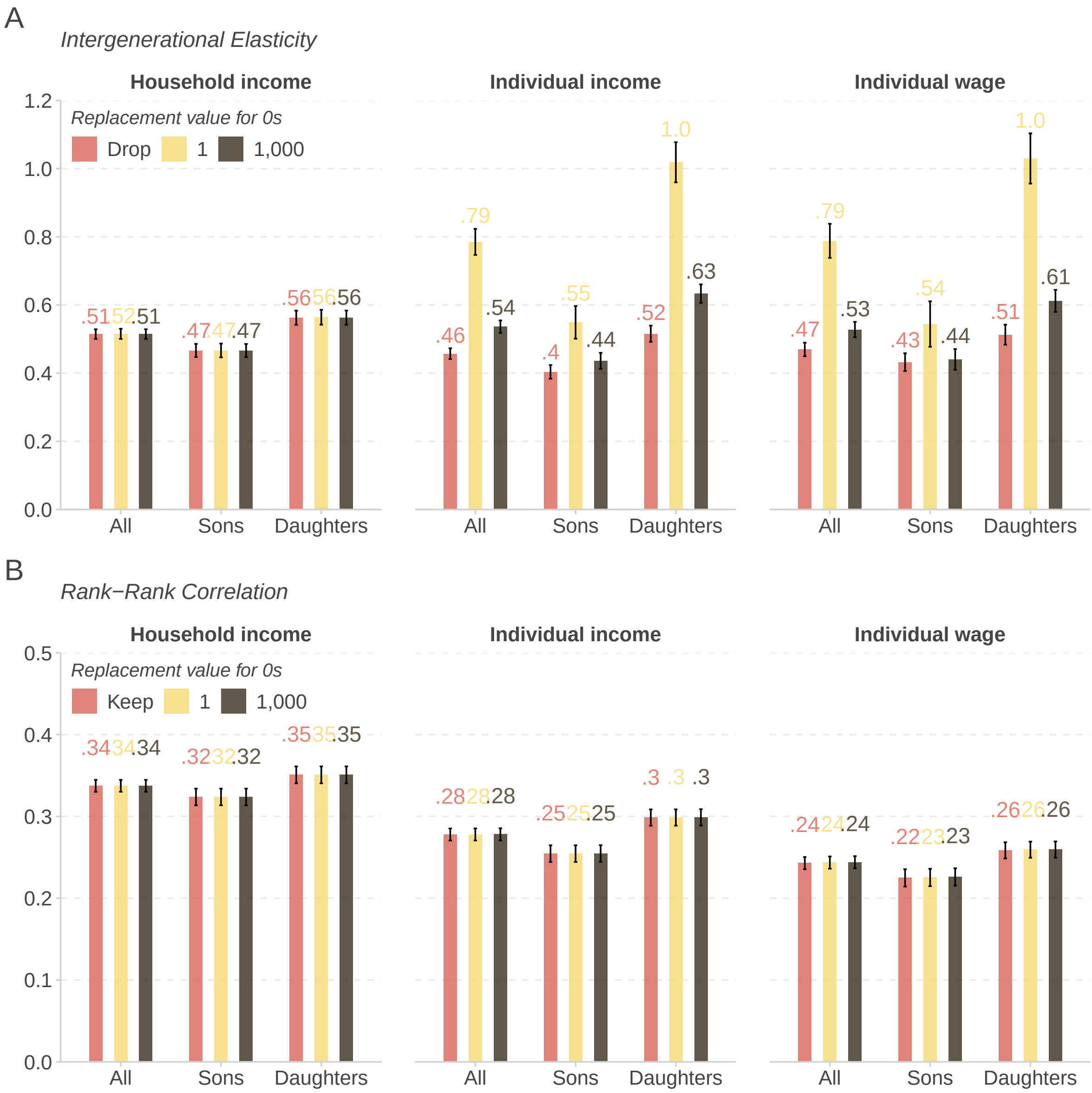
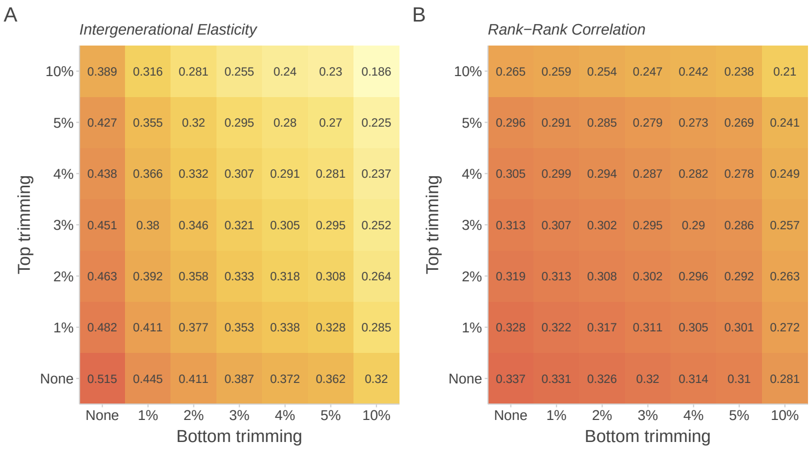
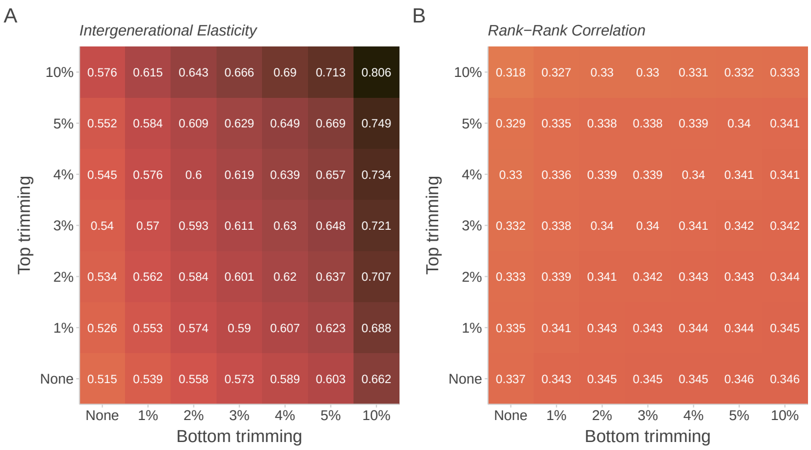



**Figure B.8:** Sensitivity to Different Zero Child Income Replacement Values

*Notes:* This figure assesses the robustness of our baseline IGE and RRC estimates to replacing incomes of children reporting negative or zero incomes by 1 euro or 1,000 euros, for different child income definitions. See Section 3 for for details on data, sample and income definitions.




**(a)** Child Income Trimming




**(b)** First-Stage Synthetic Parent Income Trimming


**(c)** Second-Stage Parent Income Trimming

**Figure B.9:** Sensitivity to Child and Parent Income Distributions Trimming

*Notes:* This figure assesses the robustness of our baseline intergenerational income mobility estimates presented to trimming the tails of the parent and child income distributions. Each cell displays the value of the corresponding intergenerational mobility measure obtained after trimming the income distribution of the corresponding sample by the fraction indicated on the x-axis at the bottom and by that indicated on the y-axis at the top. See Figure 1's notes for details on data, sample and income definitions.

# C Transition Probabilities at the Top

To analyze persistence at the top of the parent income distribution, we estimate transition matrices for the top 10%, top 5% and top 2% of parent incomes and compare our results with those from the United States.[61] We estimate the likelihood of remaining in the top 10% to be about 29% in France close to the United States figure of 26%. This statistic is almost 3 times larger than would be observed in a world where child income is unrelated to parent income (i.e., 10%). This persistence at the top gets stronger as we zoom into the top 5% (22% remaining in top 5%) and top 2% (14% remaining in top 2%). The ratio of observed persistence to counterfactual world with no link between incomes increases with parent income rank in the distribution. This suggests that mechanisms of intergenerational persistence at the top of the parent income distribution might differ from those at play for the rest of the distribution.

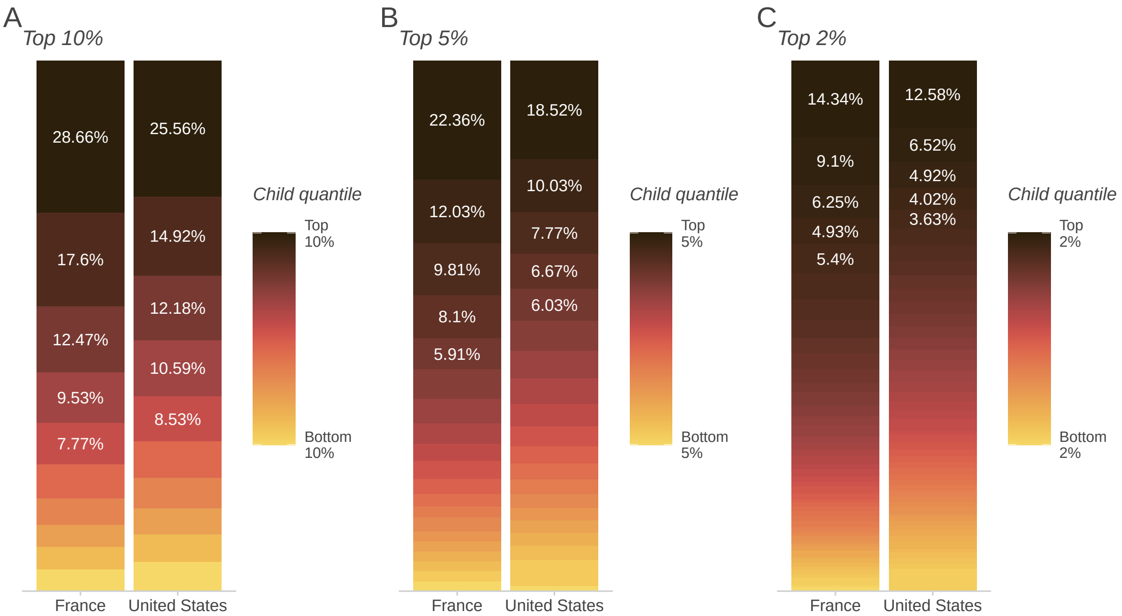



**Figure C.1:** Top Parent Income Quantiles Transition Matrices in France and United States

*Notes:* This figure presents intergenerational transition matrix estimates for children coming from families in the top 10% (panel A), top 5% (panel B) and top 2% (panel C) of the parent income distribution. We compare the transition probabilities we obtain for France with those computed by Chetty et al. (2014) for the United States. Each cell corresponds to the percentage of children in a given income quantile who have parents in a given parent income quantile. See Section 3 for for details on data, sample and income definitions.

[61] We use the detailed percentile-by-percentile estimates provided in the online appendix of Chetty et al. (2014).

# D Correlation with Local Characteristics

To pin down potential sources of the spatial variations in intergenerational mobility, we explore the department characteristics that it might correlate with. We consider an initial set of 14 variables, described in Appendix Table D.1 below, classified into 5 groups: demographic, economic, inequality, education, and social capital variables. We measure these variables as close to 1990 as possible so as to reflect the environment individuals grew up in.[62] We start by regressing department-level intergenerational mobility estimates on each of these variables in separate regressions. Both the department intergenerational mobility estimates and the characteristics are standardized, implying that the coefficients can be interpreted as correlations. Results are presented in Appendix Tables D.2 to D.4 and summarized in Appendix Figure D.1.[63]

| Variable | Definition | Source |
|---|---|---|
| **Demographic** | | |
| Density | Log number of inhabitants per square meter | 1990 BDCOM[1] |
| % Foreigner | Share without French nationality | 1990 Census |
| % Single mothers | Share of single mothers in the adult population ($\geq$ 18) | 1990 Census |
| **Economic** | | |
| Mean wage | Log average wage | 1996 DADS Panel |
| % Unemployed | Unemployment rate | 1990 Census |
| **Inequality** | | |
| Gini index | Gini index of wage inequality | 1996 DADS Panel |
| Theil index | Theil index of spatial wage segregation | 1996 DADS Panel |
| Share top 1% | Share of total wage accrued by the top 1% of wage earners | 1996 DADS Panel |
| **Education** | | |
| # HEI | Number of higher education institutions | 2007 BPE[2] |
| Distance to HEI | Average distance to the closest public higher education institution | 2007 BPE[2] |
| % HS graduates | Share of high-school graduates in adult population ($\geq$ 18) | 1990 Census |
| **Social capital** | | |
| Cultural amenities | Number of cinemas and museums per capita | 2007 BPE[2], Min. de la Culture |
| Crime | Number of offenses and crimes per capita | Min. de l'Intérieur |
| % Voters | Participation rate to the first round of the 1995 presidential election | Min. de l'Intérieur |

*Notes:*
[1] Base de données communales du recensement de la population
[2] Base permanente des équipements

**Table D.1:** Definitions and Source of Department Characteristics

There are three main take-aways. First, the IGE appears to only be significantly related to the unemployment rate, with a correlation of 0.32. This association is indeed

[62]The department-level variations of these variables are presented in Appendix Figure D.4.

[63]Note that for the IGE and RRC, a positive coefficient implies the characteristic is positively correlated with intergenerational *persistence* (i.e., negatively correlated with intergenerational *mobility*), while for absolute upward mobility a positive coefficient implies the characteristic is positively correlated with higher incomes for children born to low-income families.

striking visually when comparing the spatial distributions of the two variables (Figure 7a and Appendix Figure D.4d). Second, absolute upward mobility tends to exhibit much stronger relationships with department characteristics in general, than either the IGE or the RRC. This suggests that factors that affect absolute mobility might differ from those that affect relative mobility. Third, we find no evidence of a within France "Great Gatsby Curve". The latter refers to the positive correlation between intergenerational income persistence (defined by the IGE) and income inequality (defined by the Gini index) found across countries (Corak, 2013). The Gini index is significantly positively related to absolute upward mobility, the opposite sign one might expect if inequality is detrimental to intergenerational mobility. This contrasts with findings from Italy (Acciari, Polo and Violante, forthcoming) and North America (Chetty et al. (2014) for the United States and Corak (2020) for Canada).

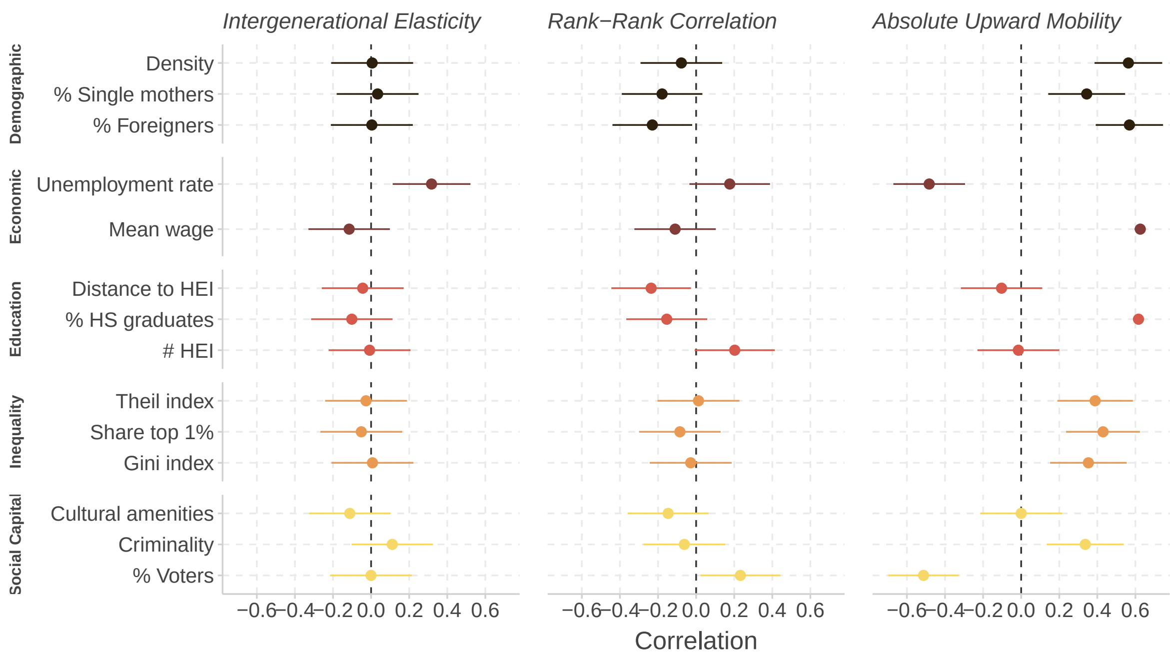



**Figure D.1:** Intergenerational Mobility and Department Characteristics - Separate Estimation

*Notes:* This figure presents the regression coefficient between department-level intergenerational mobility and department characteristics. Each coefficient is obtained from a separate regression. Both the department intergenerational mobility estimates and the characteristics are standardized, implying that the coefficients can be interpreted as correlations. See Figures 1 and 7's notes for details on data, sample and income definitions, and Appendix Table D.1 for definitions and sources of the department characteristics.

Appendix Figure D.2 provides a potential explanation to this finding by documenting the correlation between all department characteristics. The 14 variables considered are for the most part quite strongly correlated with one another, both within and between variable groups. For instance, the Gini index is highly correlated with other inequality measures, but also with population density and the share of high school graduates, two variables whose relationship with absolute upward mobility is positive. Therefore we estimate a lasso regression in order to identify the characteristics that are the most strongly associated with intergenerational mobility. The results are presented in Appendix Figure D.3.

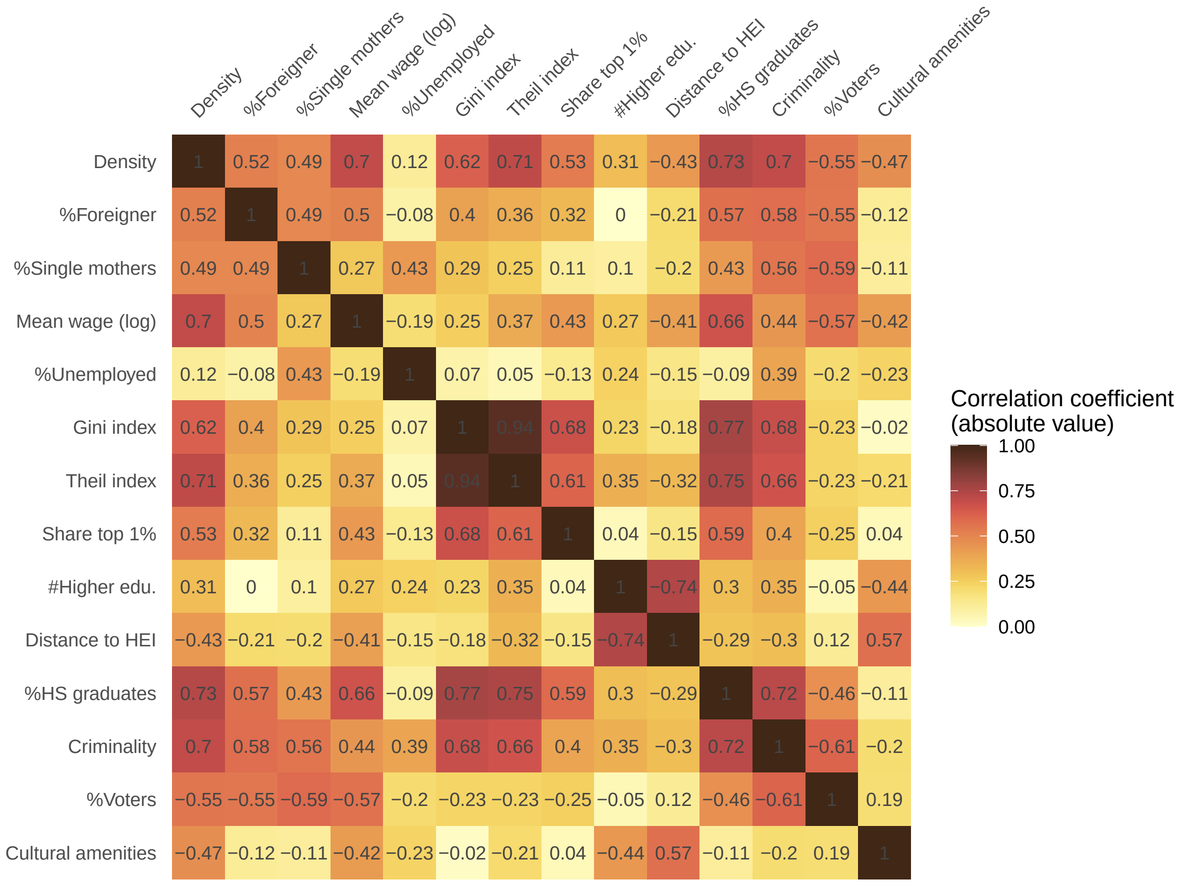



**Figure D.2:** Correlation Between Department Characteristics

*Notes:* This figure presents the correlation coefficient between all department characteristics considered, defined as follows. See Appendix Table D.1 for definitions and sources of the department characteristics.

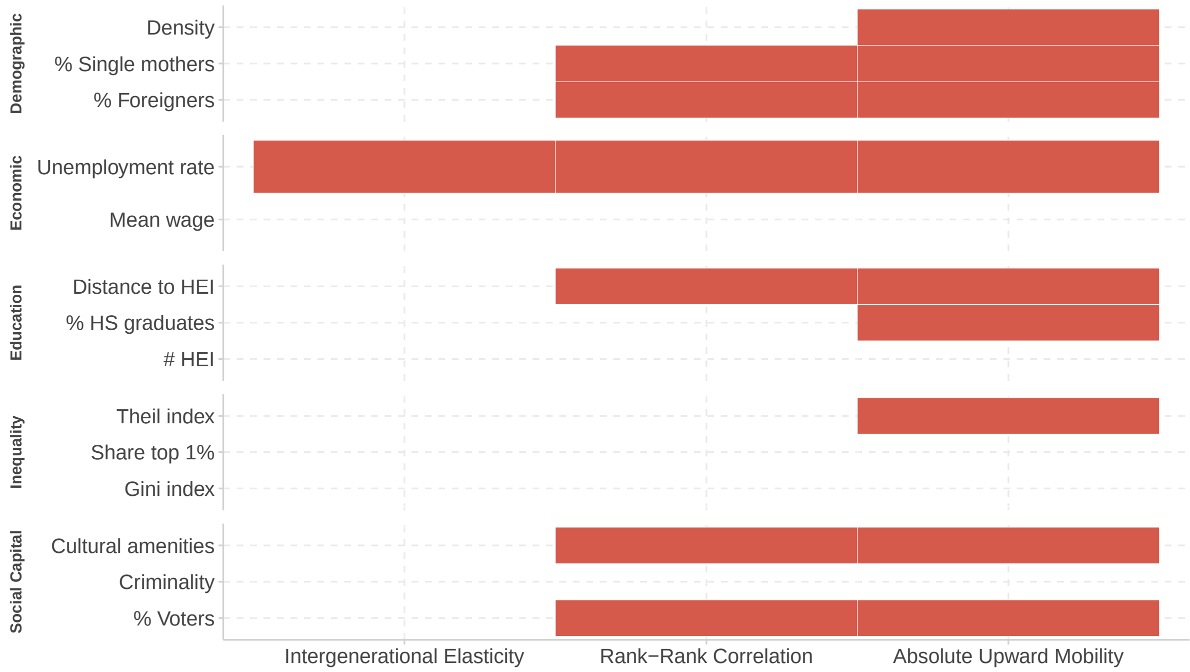



**Figure D.3:** Department Characteristics Kept by Lasso

*Notes:* This figure presents the department characteristics kept by the lasso regression. See Appendix Table D.1 for definitions and sources of the department characteristics.

The lasso analysis does not change the picture much. For the IGE, only the unemployment rate is picked up, as was the case in the univariate setting. For the RRC, the lasso maintains some demographic characteristics (% of single mothers and % foreigners), the unemployment rate, distance to higher education institution and two measures of social capital (cultural amenities and % voters). Again, these results are very much in line with what was observed in the univariate regressions. Lastly, many more characteristics are kept for absolute upward mobility.

Though the relationships we document between intergenerational mobility and department characteristics are overall pretty intuitive, these descriptive relationships cannot distinguish a potential causal effect of place from sorting. We leave this causal assessment to future studies. Still, the evidence we put forward on the potential sources of intergenerational persistence lends credence to the idea that local characteristics such as income inequality, access to cultural amenities, and labor market conditions shape individuals' intergenerational mobility prospects.

| | *Dependent variable:* Intergenerational Elasticity | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) | (11) | (12) | (13) | (14) |
| Density | 0.006 | | | | | | | | | | | | | |
| | (0.110) | | | | | | | | | | | | | |
| % Single mothers | | 0.034 | | | | | | | | | | | | |
| | | (0.110) | | | | | | | | | | | | |
| % Foreigners | | | 0.004 | | | | | | | | | | | |
| | | | (0.110) | | | | | | | | | | | |
| Unemployment rate | | | | 0.318*** | | | | | | | | | | |
| | | | | (0.104) | | | | | | | | | | |
| Mean wage | | | | | −0.115 | | | | | | | | | |
| | | | | | (0.109) | | | | | | | | | |
| Distance to HEI | | | | | | −0.044 | | | | | | | | |
| | | | | | | (0.110) | | | | | | | | |
| % HS graduates | | | | | | | −0.101 | | | | | | | |
| | | | | | | | (0.109) | | | | | | | |
| # HEI | | | | | | | | −0.008 | | | | | | |
| | | | | | | | | (0.110) | | | | | | |
| Theil index | | | | | | | | | −0.027 | | | | | |
| | | | | | | | | | (0.110) | | | | | |
| Share top 1% | | | | | | | | | | −0.051 | | | | |
| | | | | | | | | | | (0.110) | | | | |
| Gini index | | | | | | | | | | | 0.007 | | | |
| | | | | | | | | | | | (0.110) | | | |
| Cultural amenities | | | | | | | | | | | | −0.112 | | |
| | | | | | | | | | | | | (0.109) | | |
| Crime | | | | | | | | | | | | | 0.111 | |
| | | | | | | | | | | | | | (0.109) | |
| % Voters | | | | | | | | | | | | | | −0.0004 |
| | | | | | | | | | | | | | | (0.110) |
| Intercept | 0.000 | 0.000 | 0.000 | 0.000 | −0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 |
| | (0.109) | (0.109) | (0.109) | (0.103) | (0.108) | (0.109) | (0.109) | (0.109) | (0.109) | (0.109) | (0.109) | (0.108) | (0.108) | (0.109) |
| Observations | 85 | 85 | 85 | 85 | 85 | 85 | 85 | 85 | 85 | 85 | 85 | 85 | 85 | 85 |
| $R^2$ | 0.00003 | 0.001 | 0.00002 | 0.101 | 0.013 | 0.002 | 0.010 | 0.0001 | 0.001 | 0.003 | 0.0001 | 0.013 | 0.012 | 0.00000 |

*Notes:* All variables are standardized such that the regression coefficient corresponds to the correlation. See Appendix Table D.1 for variable definitions and data sources. $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$.

**Table D.2:** Correlation Between Intergenerational Elasticity and Department Characteristics

| | *Dependent variable:* Rank-Rank Correlation | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) | (11) | (12) | (13) | (14) |
| Density | −0.078 | | | | | | | | | | | | | |
| | (0.109) | | | | | | | | | | | | | |
| % Single mothers | | −0.179 | | | | | | | | | | | | |
| | | (0.108) | | | | | | | | | | | | |
| % Foreigners | | | −0.230** | | | | | | | | | | | |
| | | | (0.107) | | | | | | | | | | | |
| Unemployment rate | | | | 0.176 | | | | | | | | | | |
| | | | | (0.108) | | | | | | | | | | |
| Mean wage | | | | | −0.110 | | | | | | | | | |
| | | | | | (0.109) | | | | | | | | | |
| Distance to HEI | | | | | | −0.236** | | | | | | | | |
| | | | | | | (0.107) | | | | | | | | |
| % HS graduates | | | | | | | −0.154 | | | | | | | |
| | | | | | | | (0.108) | | | | | | | |
| # HEI | | | | | | | | 0.203* | | | | | | |
| | | | | | | | | (0.107) | | | | | | |
| Theil index | | | | | | | | | 0.012 | | | | | |
| | | | | | | | | | (0.110) | | | | | |
| Share top 1% | | | | | | | | | | −0.085 | | | | |
| | | | | | | | | | | (0.109) | | | | |
| Gini index | | | | | | | | | | | −0.028 | | | |
| | | | | | | | | | | | (0.110) | | | |
| Cultural amenities | | | | | | | | | | | | −0.147 | | |
| | | | | | | | | | | | | (0.109) | | |
| Crime | | | | | | | | | | | | | −0.062 | |
| | | | | | | | | | | | | | (0.110) | |
| % Voters | | | | | | | | | | | | | | 0.232** |
| | | | | | | | | | | | | | | (0.107) |
| Intercept | −0.000 | 0.000 | −0.000 | −0.000 | −0.000 | −0.000 | −0.000 | −0.000 | −0.000 | −0.000 | −0.000 | −0.000 | −0.000 | −0.000 |
| | (0.109) | (0.107) | (0.106) | (0.107) | (0.108) | (0.106) | (0.108) | (0.107) | (0.109) | (0.109) | (0.109) | (0.108) | (0.109) | (0.106) |
| Observations | 85 | 85 | 85 | 85 | 85 | 85 | 85 | 85 | 85 | 85 | 85 | 85 | 85 | 85 |
| $R^2$ | 0.006 | 0.032 | 0.053 | 0.031 | 0.012 | 0.056 | 0.024 | 0.041 | 0.0002 | 0.007 | 0.001 | 0.022 | 0.004 | 0.054 |

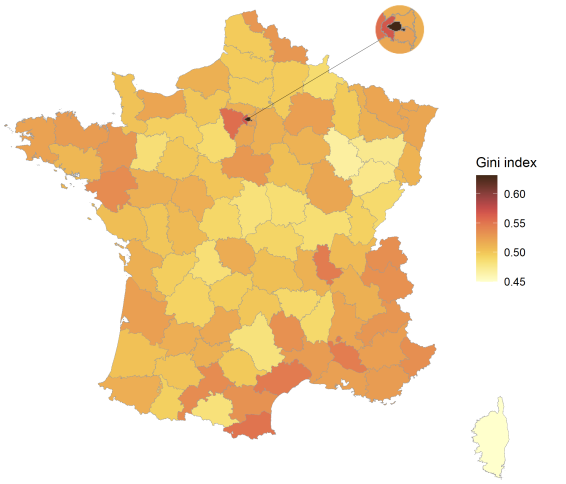
*Notes:* All variables are standardized such that the regression coefficient corresponds to the correlation. See Appendix Table D.1 for variable definitions and data sources. $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$.

**Table D.3:** Correlation Between Rank-Rank Correlation and Department Characteristics

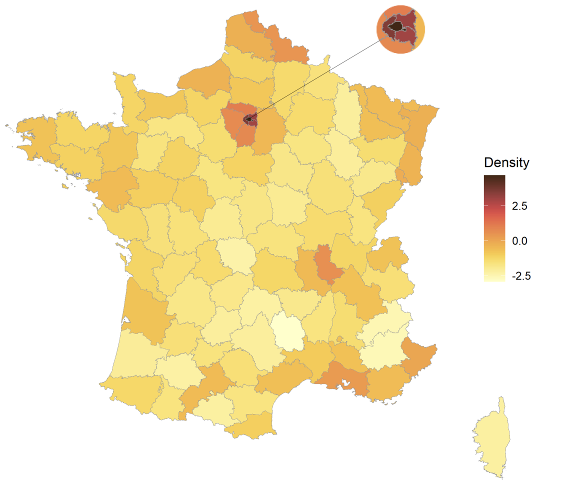
| | *Dependent variable:* Absolute Upward Mobility | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) | (11) | (12) | (13) | (14) |
| Density | $0.563^{***}$ | | | | | | | | | | | | | |
| | (0.091) | | | | | | | | | | | | | |
| % Single mothers | | $0.344^{***}$ | | | | | | | | | | | | |
| | | (0.103) | | | | | | | | | | | | |
| % Foreigners | | | $0.568^{***}$ | | | | | | | | | | | |
| | | | (0.090) | | | | | | | | | | | |
| Unemployment rate | | | | $-0.483^{***}$ | | | | | | | | | | |
| | | | | (0.096) | | | | | | | | | | |
| Mean wage | | | | | $0.626^{***}$ | | | | | | | | | |
| | | | | | (0.086) | | | | | | | | | |
| Distance to HEI | | | | | | −0.103 | | | | | | | | |
| | | | | | | (0.109) | | | | | | | | |
| % HS graduates | | | | | | | $0.616^{***}$ | | | | | | | |
| | | | | | | | (0.086) | | | | | | | |
| # HEI | | | | | | | | −0.015 | | | | | | |
| | | | | | | | | (0.110) | | | | | | |
| Theil index | | | | | | | | | $0.388^{***}$ | | | | | |
| | | | | | | | | | (0.101) | | | | | |
| Share top 1% | | | | | | | | | | $0.430^{***}$ | | | | |
| | | | | | | | | | | (0.099) | | | | |
| Gini index | | | | | | | | | | | $0.353^{***}$ | | | |
| | | | | | | | | | | | (0.103) | | | |
| Cultural amenities | | | | | | | | | | | | 0.0001 | | |
| | | | | | | | | | | | | (0.110) | | |
| Crime | | | | | | | | | | | | | $0.337^{***}$ | |
| | | | | | | | | | | | | | (0.103) | |
| % Voters | | | | | | | | | | | | | | $-0.513^{***}$ |
| | | | | | | | | | | | | | | (0.094) |
| Intercept | −0.000 | −0.000 | −0.000 | −0.000 | 0.000 | −0.000 | −0.000 | −0.000 | −0.000 | −0.000 | −0.000 | −0.000 | −0.000 | −0.000 |
| | (0.090) | (0.102) | (0.090) | (0.096) | (0.085) | (0.109) | (0.086) | (0.109) | (0.101) | (0.099) | (0.102) | (0.109) | (0.103) | (0.094) |
| Observations | 85 | 85 | 85 | 85 | 85 | 85 | 85 | 85 | 85 | 85 | 85 | 85 | 85 | 85 |
| $R^2$ | 0.317 | 0.119 | 0.323 | 0.233 | 0.391 | 0.011 | 0.380 | 0.0002 | 0.151 | 0.185 | 0.125 | 0.00000 | 0.113 | 0.263 |

*Notes:* All variables are standardized such that the regression coefficient corresponds to the correlation. See Appendix Table D.1 for variable definitions and data sources. $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$.

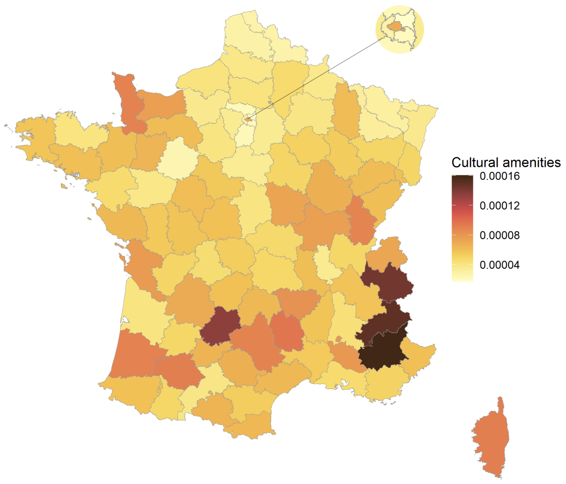
**Table D.4:** Correlation Between Absolute Upward Mobility and Department Characteristics



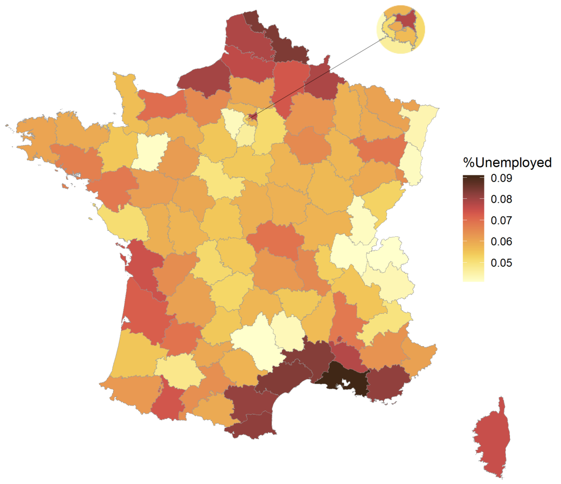

**(a)** Gini Index of Inequality



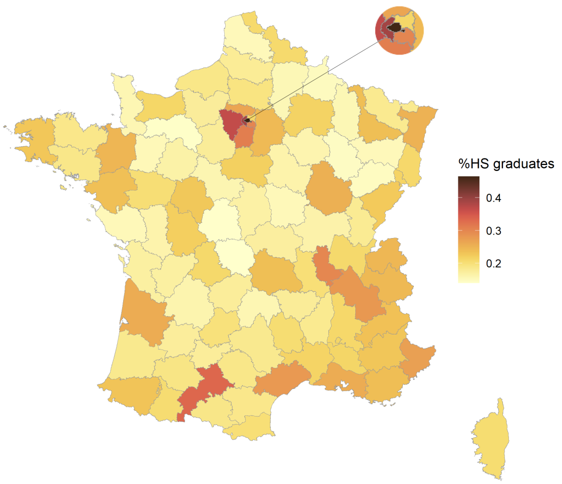

**(b)** Log Population Density




**(c)** Museums and Cinemas per Capita




**(d)** Unemployment Rate




**(e)** Share of High School Graduates

**Figure D.4:** Spatial Distribution of Average Department Characteristics

*Notes:* This Figure presents the spatial variations of the department characteristics we correlate with our intergenerational mobility estimates to produce Figure D.1. The definition and data source of each of these department characteristics are detailed in Table D.1.

# E Appendix Figures

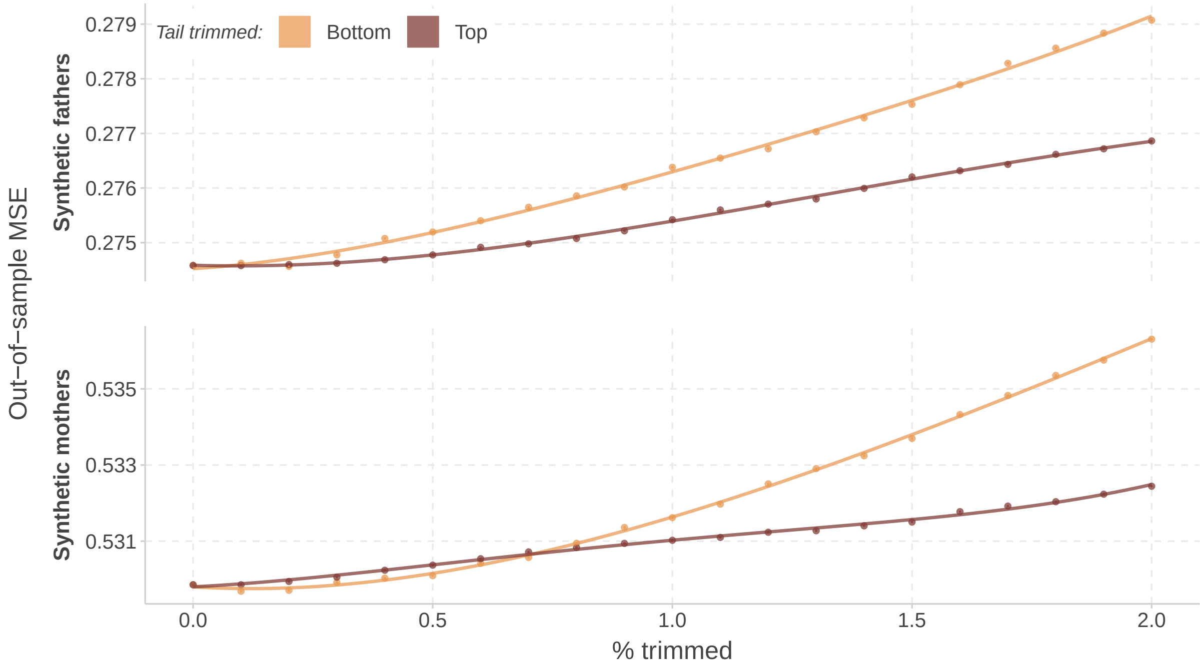



**Figure E.1:** Out-of-sample MSE as a Function of Top and Bottom Trimming

*Notes:* This figure plots the out-of-sample MSE as a function of trimming various shares of the tails of synthetic parents' income distribution. Our-of-sample MSEs correspond to the average MSE obtained from 5-fold cross-validation. See Sections 3.1 and 3.2 for details on the exact model being estimates and sample construction.

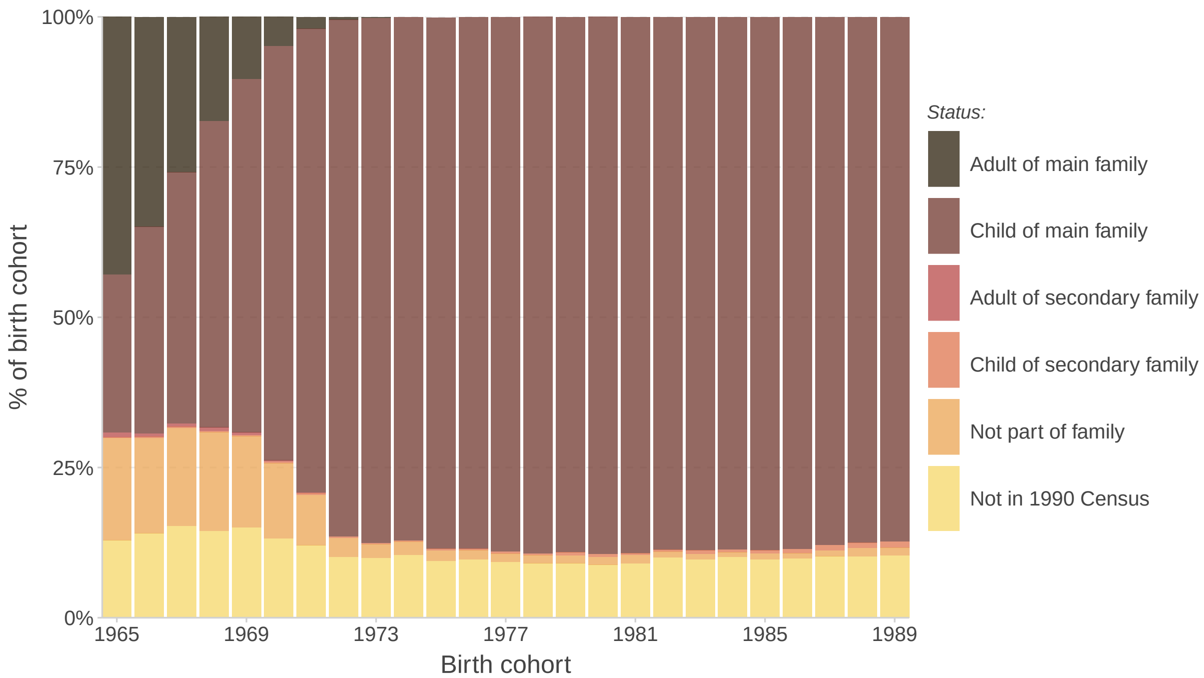



**Figure E.2:** Family Position in 1990 Census by Child Birth Cohort

*Notes:* This figure presents the family position of EDP individuals in the 1990 census by birth cohort. The sample is restricted to EDP individuals born in metropolitan France.

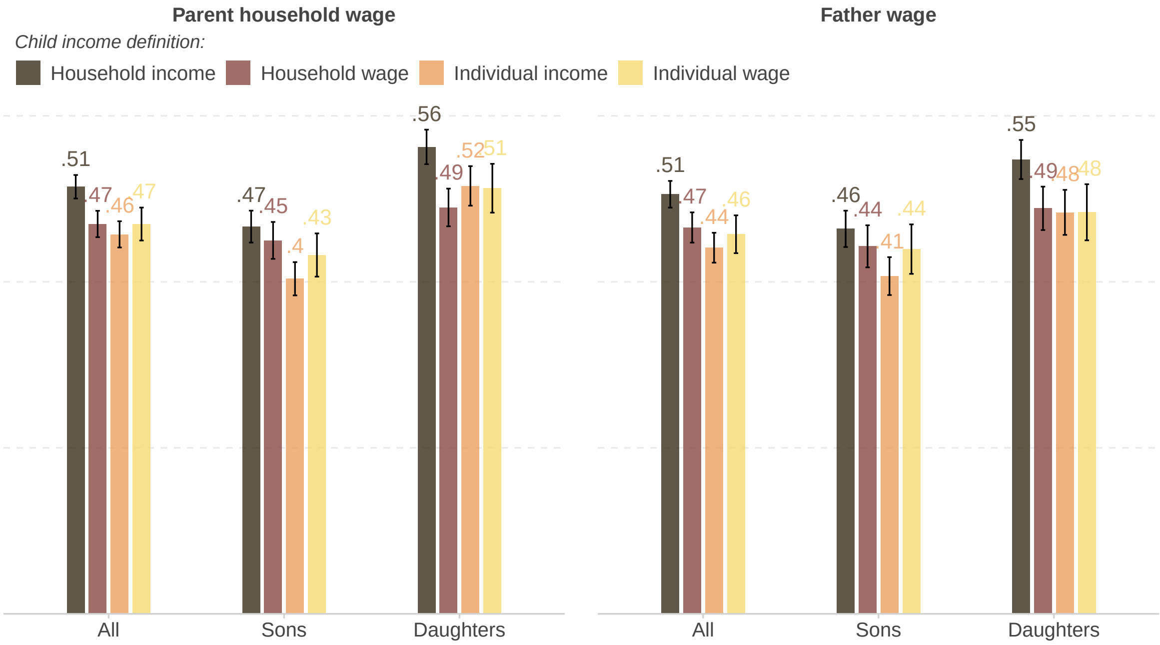



**Figure E.3:** Baseline IGE Estimates for Different Child and Parent Income Definitions

*Notes:* This figure presents our baseline intergenerational income elasticity estimates for various parent and child income definitions. Each bar represents the coefficient of an OLS regression of child income on parent income, for the entire sample (All) and for sons and daughters separately. See Section 3 for for details on data, sample and income definitions.

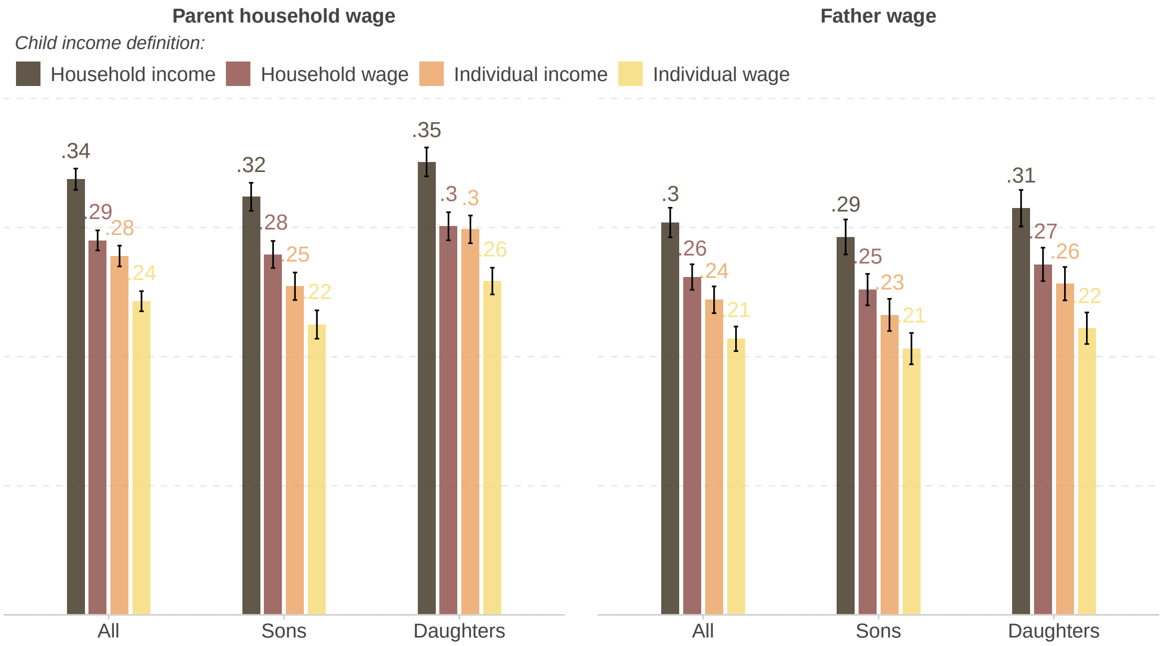



**Figure E.4:** Baseline RRC Estimates for Different Child and Parent Income Definitions

*Notes:* This figure presents our baseline intergenerational rank-rank correlation estimates for various parent and child income definitions. Each bar represents the coefficient of an OLS regression of child income rank on parent income rank, for the entire sample (All) and for sons and daughters separately. See Section 3 for for details on data, sample and income definitions.

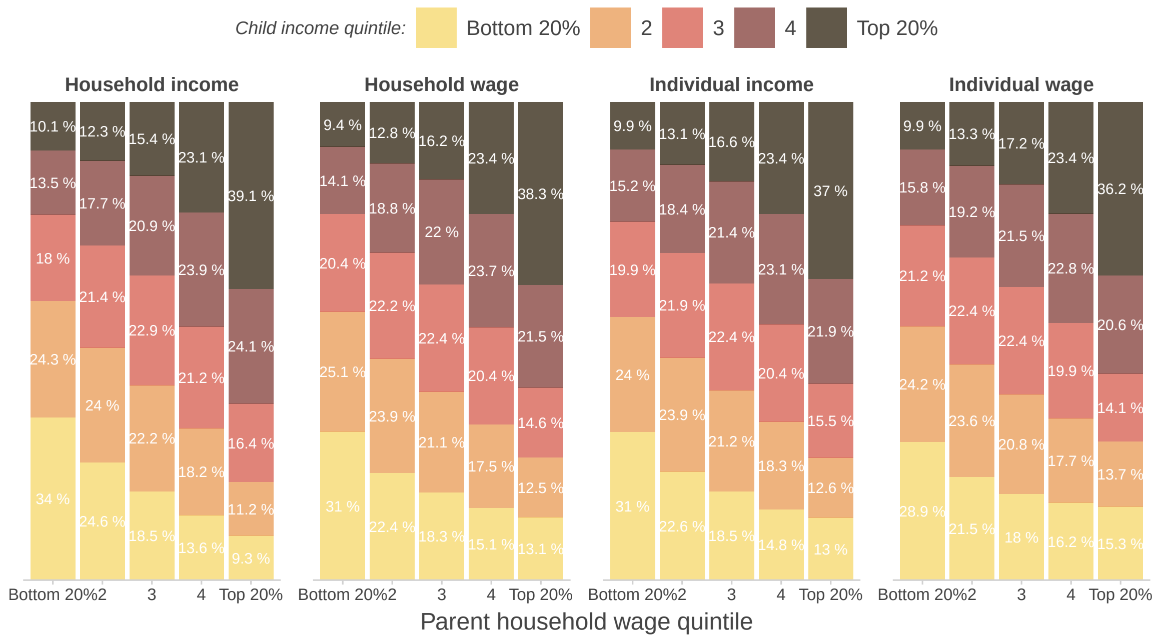



**Figure E.5:** Baseline Quintile Transition Matrix for Different Child Income Definitions

*Notes:* This figure presents our baseline intergenerational transition matrix estimates for various child income definitions. Each cell corresponds to the percentage of children in a given income quintile who have parents in a given parent income quintile. See Section 3 for for details on data, sample and income definitions.

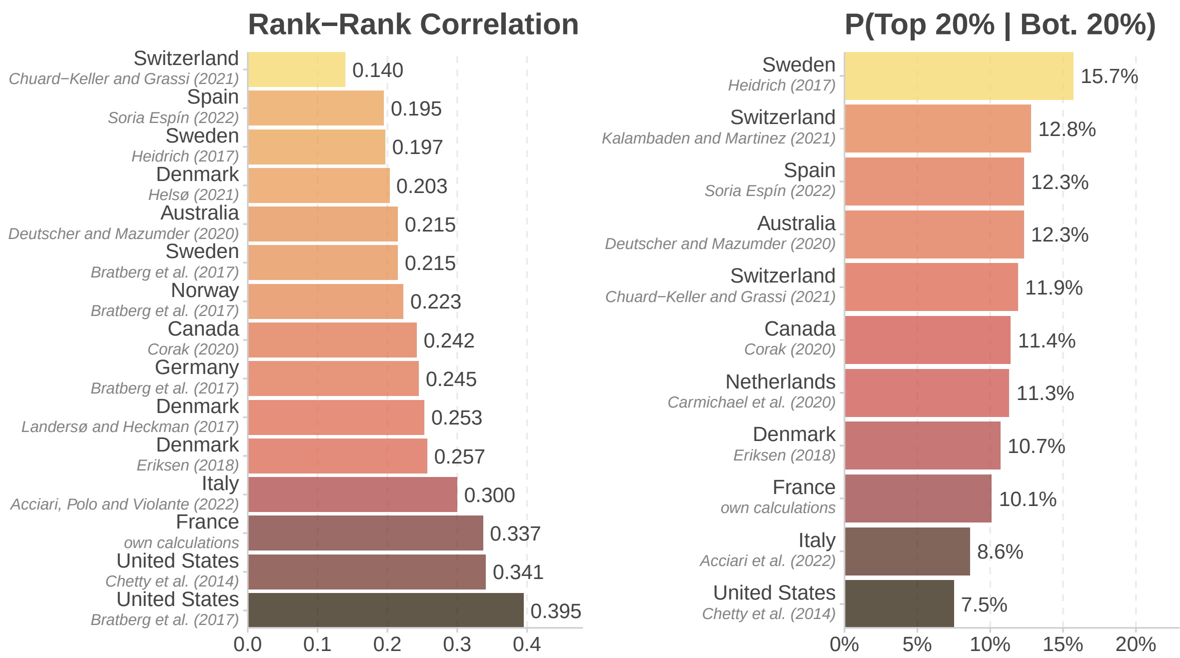



**Figure E.6:** Rank-Rank Correlation and Upward Mobility in International Comparison

*Notes:* This figure represents the international comparisons in rank-rank correlation and transition matrix cells presented in Tables 1 and 2.

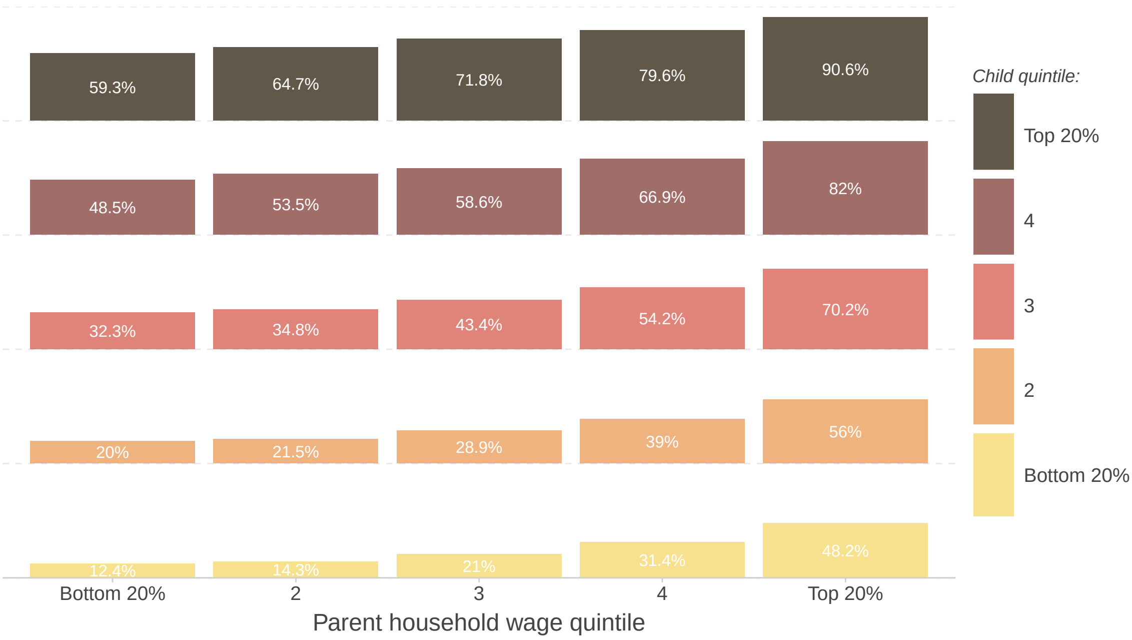



**Figure E.7:** Higher Education Graduation by Quintile Transition Matrix Cell

*Notes:* This figure presents the percentage of children graduating from higher education in each cell of the quintile transition matrix. Each cell corresponds to the percentage of children in a given income quintile coming from a family in a given parent income quintile who have a higher education diploma. See Sections 3 and 4.4 for for details on data, sample and income definitions.


**Figure E.8:** French Departments

*Notes:* This figure represents the 96 metropolitan French departments. The borders of these departments have not changed over the study period. For convenience, we treat Corsica (*Haute Corse* and *Corse du Sude*) as a single department.


**Figure E.9:** Illustration of Absolute Upward Mobility for the *Nord* Department

*Notes:* This figure presents a non-parametric binned scatter plot of the relationship between child income rank and parent income rank for individuals who grew up in the *Nord* department. The dashed line shows the expected income rank, here 37.4, for children whose parents locate at the 25$^{th}$ percentile. The orange line is a linear regression fit through the conditional expectation. See Figure 1's notes for details on data, sample and income definitions.

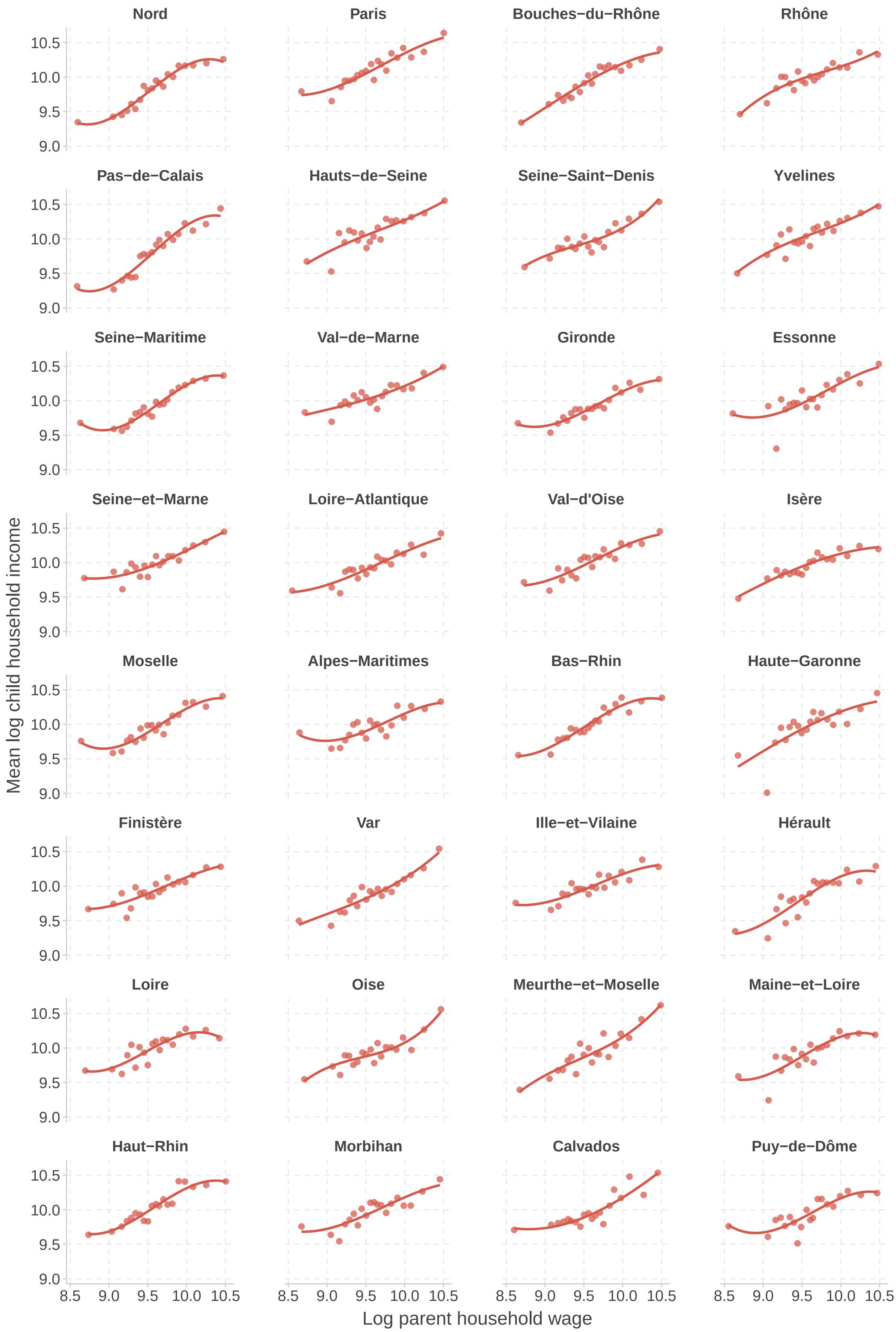



**Figure E.10:** Department-Level Log-Log Relationships

*Notes:* This figure presents the non-parametric binned scatter plot of the relationship between child log income and parent log income separately for each childhood department. The childhood department is that observed in 1990, i.e., when individuals were between 9 and 18 years old. See Figure 1's notes for details on data, sample and income definitions.

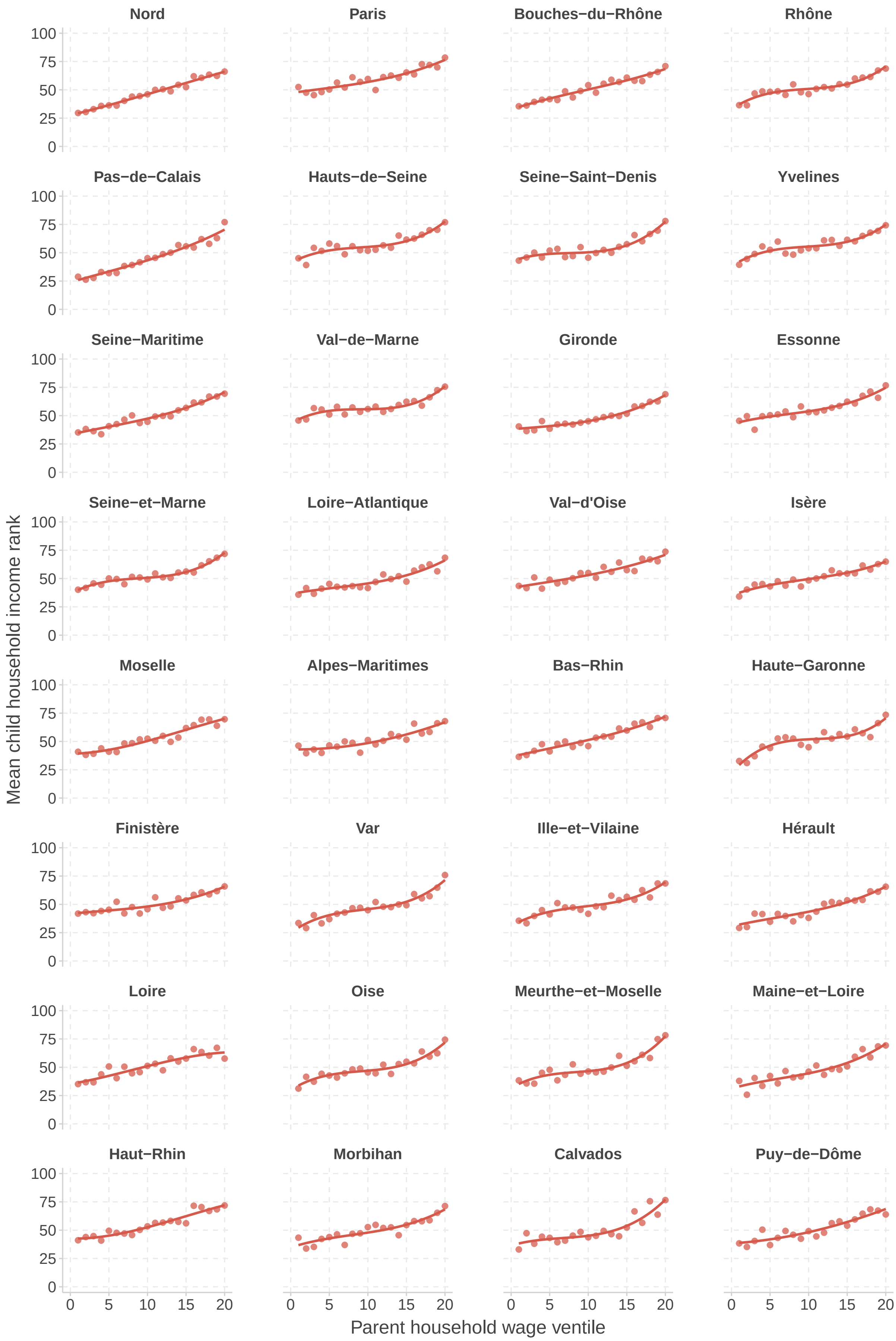



**Figure E.11:** Department-Level Rank-Rank Relationships

*Notes:* This figure presents the non-parametric binned scatter plot of the relationship between child income rank and parent income rank separately for each childhood department. The childhood department is that observed in 1990, i.e., when individuals were between 9 and 18 years old. See Figure 1's notes for details on data, sample and income definitions.

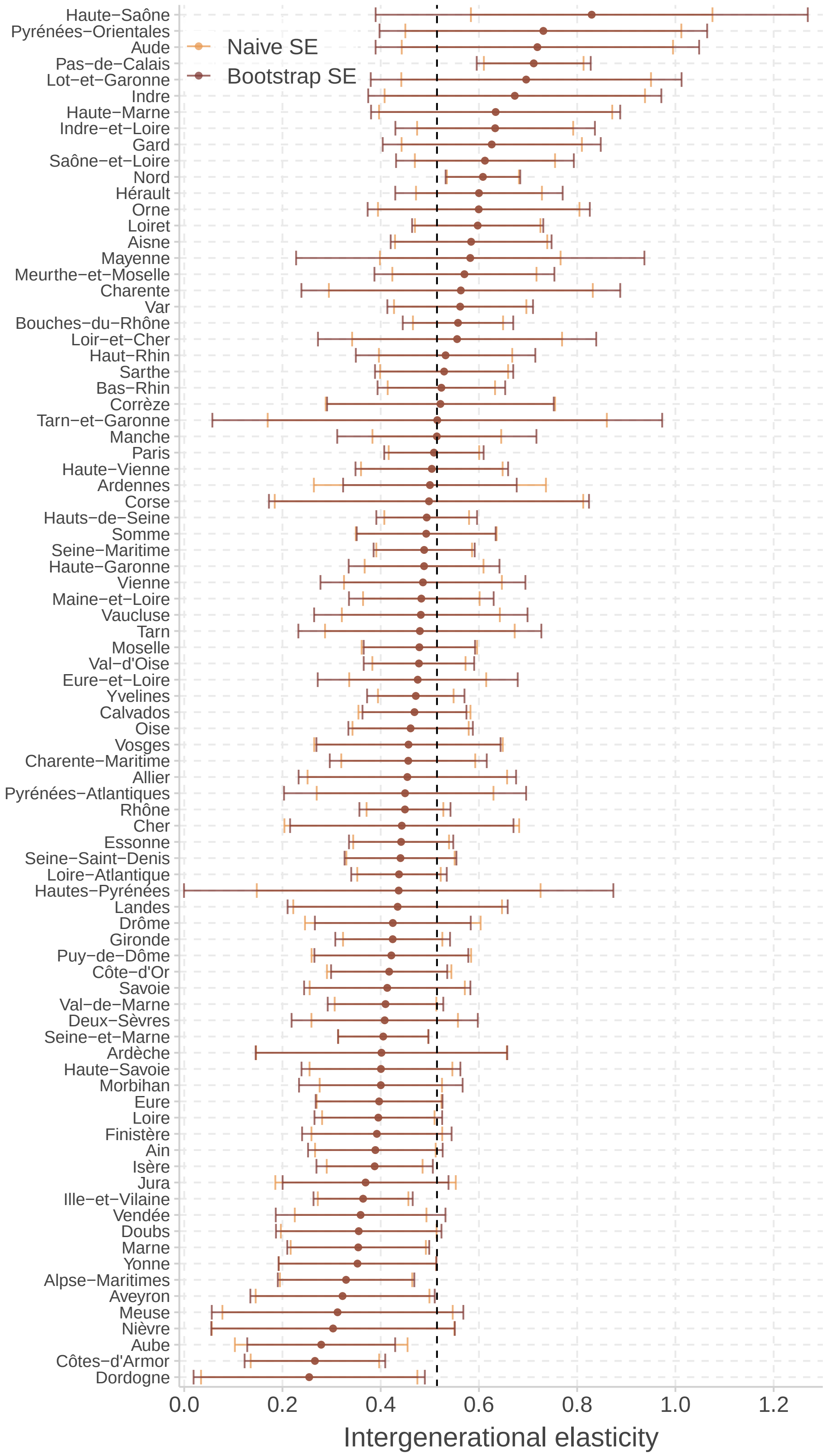



**Figure E.12:** Department-Level Intergenerational Elasticities

*Notes:* This figure presents the intergenerational elasticity in household income and its confidence interval, estimated separately for each childhood department with more than 200 observations. The childhood department is that observed in 1990, i.e., when individuals were between 9 and 18 years old. The dashed black line represents the national estimate. See Figure 1's notes for details on data, sample and income definitions.

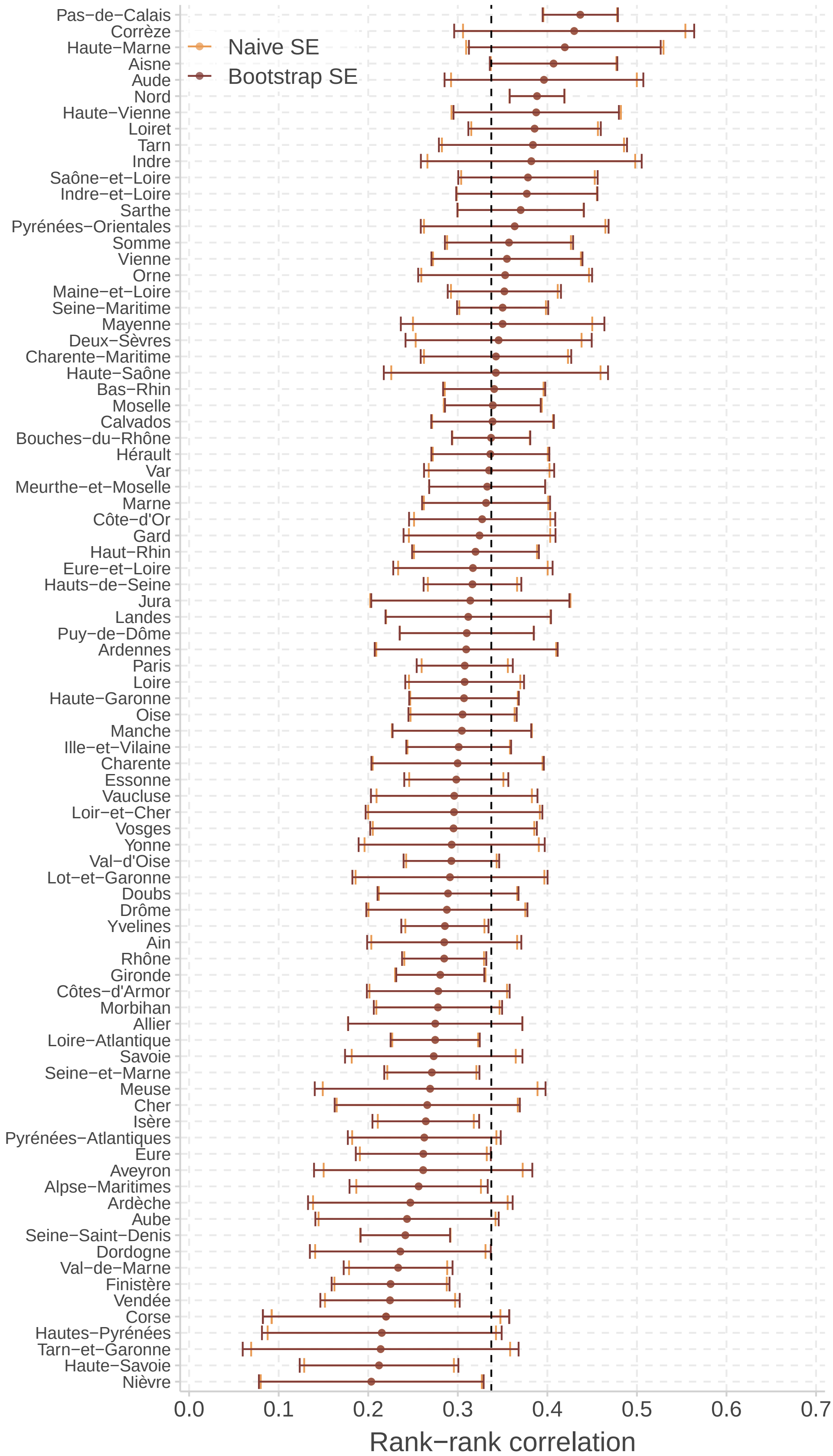



**Figure E.13:** Department-Level Rank-Rank Correlations

*Notes:* This figure presents the rank-rank correlation in household income and its confidence interval, estimated separately for each childhood department with more than 200 observations. The childhood department is that observed in 1990, i.e., when individuals were between 9 and 18 years old. The dashed black line represents the national estimate. See Figure 1's notes for details on data, sample and income definitions.


**Figure E.14:** Geographic Mobility by Parent Household Wage Rank

*Notes:* This figure presents the percentage of movers by parent income rank. Movers are defined as individuals whose adulthood department of residence is different from that of their childhood. See Figure 1 and 8's notes for details on data, sample and income definitions.

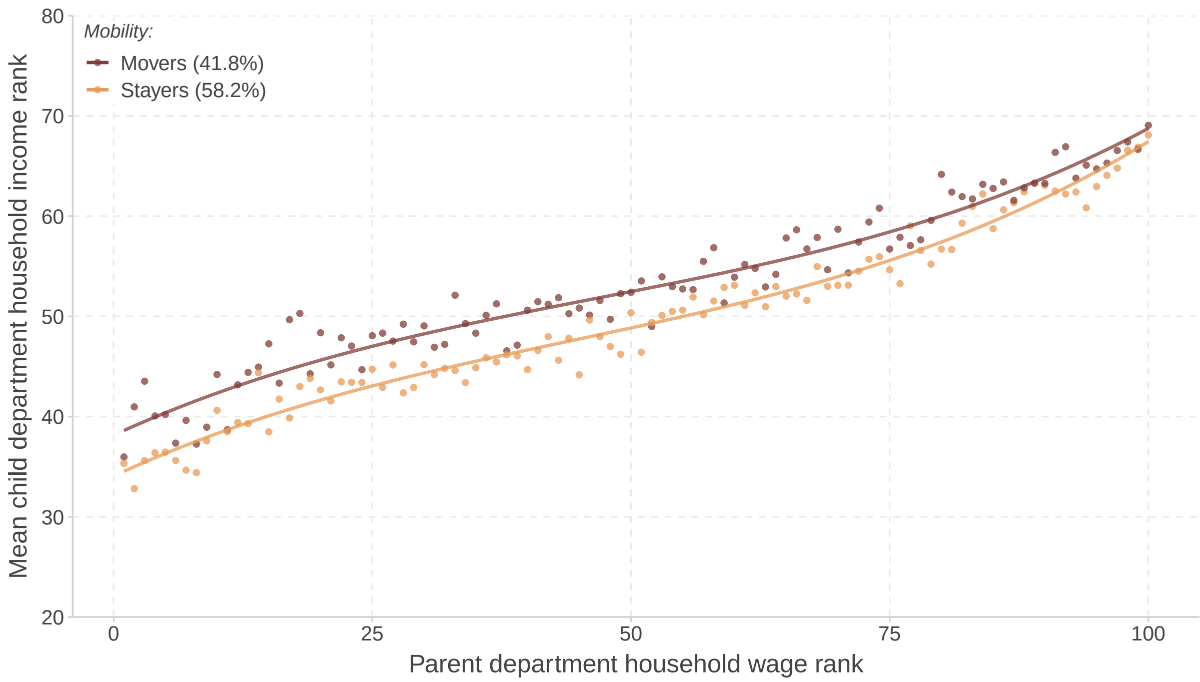



**Figure E.15:** Intergenerational Mobility and Geographic Mobility - Department Ranks

*Notes:* This figure represents the conditional expectation function of child household income rank with respect to parent household wage rank separately for individuals whose adulthood department of residence is different or not from their childhood department of residence. Percentile ranks are computed according to the local department income distribution. Parents are ranked within their department of residence in 1990 while children are ranked within their adulthood department. See Figures 1 and 8's notes for details on data, sample and income definitions.

# F Appendix Tables

| Birth Cohort | Born in Metropolitan France | + Live with parents in 1990 census | + At least one obs. in tax returns data | + At least one obs. in tax returns data at 35-45 | + No parent in occupation 1 or 31 |
|---|---|---|---|---|---|
| 1972 | 9,083 | 7,946 | 7,582 | 7,582 | 7,077 |
| 1973 | 8,647 | 7,670 | 7,330 | 7,330 | 6,788 |
| 1974 | 8,704 | 7,713 | 7,372 | 7,372 | 6,830 |
| 1975 | 7,334 | 6,565 | 6,290 | 6,290 | 5,873 |
| 1976 | 7,762 | 6,963 | 6,662 | 6,650 | 6,199 |
| 1977 | 7,972 | 7,175 | 6,886 | 6,848 | 6,395 |
| 1978 | 7,755 | 7,000 | 6,748 | 6,677 | 6,224 |
| 1979 | 8,473 | 7,620 | 7,351 | 7,233 | 6,770 |
| 1980 | 8,822 | 7,965 | 7,642 | 7,426 | 6,945 |
| 1981 | 8,457 | 7,631 | 7,344 | 6,958 | 6,531 |
| 1972-1981 | 83,009 | 74,248 | 71,207 | 70,366 | 65,632 |

**Table F.1:** Child Sample Construction

| Characteristic | Synthetic Parents | Actual Parents |
|---|---|---|
| Females | 53.42% | 52.26% |
| Age in 1990 | 41.22 | 40.74 |
| Born French | 89.95% | 88.36% |
| *1-digit occupation* | | |
| 1. Farmers | 3.72% | 3.47% |
| 2. Craftsmen, salespeople, and heads of businesses | 6.98% | 6.77% |
| 3. Managerial and professional occupations | 9.76% | 9.35% |
| 4. Intermediate professions | 15.48% | 15.35% |
| 5. Employees | 20.76% | 20.39% |
| 6. Blue collar workers | 23.19% | 24.6% |
| 7. Retirees | 1.30% | 1.32% |
| 8. Other with no professional activity | 18.81% | 18.76% |
| *Education* | | |
| No diploma | 22.45% | 23.80% |
| Primary education | 19.38% | 18.93% |
| BEPC | 7.99% | 8.18% |
| CAP | 20.76% | 19.91% |
| BEP | 4.95% | 5.00% |
| High school diploma | 11.64% | 11.47% |
| Bachelor or technical degree | 6.08 | 6.18% |
| Masters or PhD | 6.75% | 6.52% |
| *Country of birth* | | |
| France | 86.18% | 84.81% |
| Maghreb | 6.62% | 8.03% |
| Other Africa | 0.55% | 0.73% |
| South Europe | 3.32% | 3.33% |
| Other Europe | 2.33% | 2.17% |
| Rest of the world | 1.00% | 0.94% |
| *Family structure* | | |
| Single fathers | 0.93% | 0.72% |
| Single mothers | 5.58% | 5.25% |
| Both spouses active | 58.73% | 58.28% |
| Mother inactive | 31.35% | 32.32% |
| Father inactive | 1.38% | 1.38% |
| Both spouses inactive | 2.03% | 2.06% |
| *Municipality characteristics* | | |
| Log population | 782.64 | 785.50 |
| Log density | 46.42 | 49.12 |
| % foreigners | 2.31% | 2.33% |
| Unemployment rate | 6.22% | 6.26% |
| % single mothers | 6.36% | 6.40% |
| N | 134,572 | 140,136 |

*Notes:* See Section 3.2 for details on construction of samples. These statistics are computed before applying any income observation restrictions.

**Table F.2:** Average Characteristics of Actual and Synthetic Parents

| 2-digit occupation | Synthetic Parents | Actual Parents |
|---|---|---|
| Farmers with small farm | 0.92% | 0.84% |
| Farmers with medium-sized farm | 1.22% | 1.19% |
| Farmers with large farm | 1.58% | 1.44% |
| Craftsmen | 3.62% | 3.57% |
| Trade workers and related | 2.62% | 2.50% |
| Heads of company with $\geq$ 10 employees | 0.73% | 0.70% |
| Liberal profession | 1.38% | 1.32% |
| Public sector executives | 1.07% | 1.05% |
| Professors and scientific professions | 2.12% | 1.97% |
| Information, arts, and entertainment professions | 0.32% | 0.31% |
| Administrative executives and sales representatives | 2.72% | 2.66% |
| Engineers, technical executives | 2.16% | 2.05% |
| Teachers and related | 2.64% | 2.57% |
| Intermediate health and social work professions | 2.48% | 2.62% |
| Clerk, religious | 0.01% | 0.01% |
| Intermediate administrative professions of the public sector | 1.54% | 1.41% |
| Intermediate administrative professions and salesmen | 4.06% | 4.03% |
| Technicians | 2.30% | 2.29% |
| Foremen, supervisors | 2.44% | 2.42% |
| Civil servants | 6.74% | 6.69% |
| Police and military officers | 1.27% | 1.35% |
| Company administrative employees | 6.92% | 6.70% |
| Trade employees | 2.24% | 2.16% |
| Personal service workers | 3.58% | 3.49% |
| Skilled industrial workers | 5.82% | 6.14% |
| Skilled crafts workers | 4.60% | 4.83% |
| Drivers | 2.19% | 2.39% |
| Skilled handling, storing and transport workers | 1.41% | 1.47% |
| Unskilled industrial workers | 6.19% | 6.67% |
| Unskilled crafts workers | 2.32% | 2.42% |
| Agricultural workers | 0.66% | 0.69% |
| Former farmers | 0.09% | 0.07% |
| Former craftsmen, salespeople, and heads of businesses | 0.10% | 0.08% |
| Former managerial and professional occupation | 0.09% | 0.10% |
| Former intermediate professions | 0.19% | 0.17% |
| Former employees | 0.33% | 0.30% |
| Former blue collar workers | 0.51% | 0.60% |
| Unemployed who have never worked | 0.36% | 0.38% |
| Military contingent | 0.00% | 0.00% |
| Students $\geq$ 15 yrs old | 0.10% | 0.04% |
| Other inactive $\leq$ 60 yrs old | 18.24% | 18.20% |
| Other inactive $\geq$ 60 yrs old | 0.10% | 0.12% |
| N | 134,572 | 140,136 |

*Notes:* See Table F.2's notes for sample construction.

**Table F.3:** Share of Actual and Synthetic Parents by 2-Digit Occupation

| Gender | At least one child born in Metrop. France 1972-1981 | + Observed in 1990 Census | + Born even year | + At least two obs. at 35-45 in All Employee Panel | + Not in occupation 1 or 31 |
|---|---|---|---|---|---|
| Fathers | 49,746 | 43,851 | 22,227 | 16,699 | 16,450 |
| Mothers | 52,904 | 48,097 | 24,297 | 15,104 | 14,973 |
| All | 102,650 | 91,948 | 46,524 | 31,803 | 31,423 |

**Table F.4:** Synthetic Parents Sample Construction

| Child income definition | Parent income definition | Number of observations | Negative or 0 child incomes |
|---|---|---|---|
| Household income | Family income | 64,572 | 42 |
| Household income | Father income | 58,435 | 37 |
| Household wage | Family income | 64,602 | 1,982 |
| Household wage | Father income | 58,461 | 1,716 |
| Individual income | Family income | 65,609 | 2,870 |
| Individual income | Father income | 59,355 | 2,525 |
| Labor income | Family income | 65,609 | 5,385 |
| Labor income | Father income | 59,355 | 4,792 |

*Notes:* The very slight discrepancy in the number of child income observations compared to those reported in Section 3.2 comes from the fact we code to missing 23 father ages in 1990 which were below 14 and above 70.

**Table F.5:** Number of Observations by Child and Parent Income Definitions

| | N | Missing (%) | Mean | Std. Dev. | 25th pctile | Median | 75th pctile |
|---|---|---|---|---|---|---|---|
| **Synthetic Parents** | | | | | | | |
| Synthetic father income (35-45 yrs old) | 16,450 | 0 | 25,902 | 17,265 | 16,251 | 21,966 | 30,427 |
| Number of syn. father income observations | 16,450 | 0 | 7.66 | 2.42 | 6 | 8 | 9 |
| Synthetic mother income (35-45 yrs old) | 14,973 | 0 | 15,167 | 10,143 | 7,496 | 14,140 | 21,027 |
| Number of syn. mother income observations | 14,973 | 0 | 6.95 | 2.84 | 5 | 7 | 9 |
| **Parents** | | | | | | | |
| Fraction single parents in 1990 | 11.64% | | | | | | |
| Fraction female among single parents | 88.44% | | | | | | |
| Father age at child's birth | 65,632 | 0.06 | 28.7 | 5.77 | 25 | 28 | 32 |
| Mother age at child's birth | 65,632 | 0.02 | 25.91 | 5.01 | 22 | 25 | 29 |
| Father age in 1990 | 65,632 | 0.1 | 41.97 | 6.61 | 38 | 41 | 45 |
| Mother age in 1990 | 65,632 | 0.01 | 39.41 | 5.81 | 35 | 39 | 43 |
| **Children** | | | | | | | |
| Household income (average 2010-16) | 65,632 | 0.01 | 25,655 | 19,402 | 17,049 | 22,772 | 30,173 |
| Household wage (average 2010-16) | 65,632 | 0.02 | 21,426 | 15,819 | 12,761 | 20,019 | 27,069 |
| Individual income (average 2010-16) | 65,632 | 0 | 20,410 | 17,855 | 10,012 | 19,227 | 26,662 |
| Labor income (average 2010-16) | 65,632 | 0 | 22,726 | 19,005 | 13,951 | 20,593 | 27,865 |
| Fraction female | 49.59% | | | | | | |

*Notes:* See Sections 3.2 and 3.3 for details on sample construction and income definitions.

**Table F.6:** Descriptive Statistics

| | Intergenerational Elasticity | First-Stage Instruments | Data | Income Definitions | Child Age |
|---|---|---|---|---|---|
| Lefranc and Trannoy (2005) | 0.4-0.438[1] | Education (8 cat.) + occupation (7 cat.) | FQP | labor earnings (excl. UI)[2] | 30-40 |
| Lefranc (2018) | 0.577[3] | Education (6 cat.) | FQP | labor earnings (excl. UI)[2] | 28-32 |
| EqualChances.org | 0.357 | Education (3 cat.) + occupation (9 cat.) | Synthetic fathers: ECHP Sons: EU-SILC | net personal employee income | - |
| **Our estimate** | **0.439** | | | | |

*Notes:* FQP = Formation-Qualification-Profession; ECHP = European Community Household Panel; EU-SILC = European Union Statistics on Income and Living Conditions
[1] Estimates taken from Table I, Panel A, cols. (1)-(4), p.65.
[2] Only salaried workers.
[3] Estimates taken from Table 2, 1971-75, col. (2), p.823.

**Table F.7:** Comparison with Existing Father-Son IGE Estimates for France

| | IGE | RRC | AUM |
|---|---|---|---|
| First-stage MSE | $-0.005$ | $-0.025$ | $-1.401$ |
| | (0.068) | (0.033) | (2.400) |
| Constant | $0.482^{***}$ | $0.321^{***}$ | $42.737^{***}$ |
| | (0.034) | (0.016) | (1.201) |
| Observations | 85 | 85 | 85 |

*Notes:* $^{*}$p<0.1; $^{**}$p<0.05; $^{***}$p<0.01.

**Table F.8:** Department-Level MSEs and Measures of Intergenerational Income Mobility

**Table F.9:** Department-Level Intergenerational Mobility Estimates

| | Department | Observations | IGE | RRC | AUM |
|---|---|---|---|---|---|
| 01 | Ain | 535 | 0.39 | 0.28 | 46.4 |
| 02 | Aisne | 735 | 0.58 | 0.41 | 37.1 |
| 03 | Allier | 365 | 0.45 | 0.27 | 41.6 |
| 04 | Alpes-de-Haute-Provence | 141 | * | * | * |
| 05 | Hautes-Alpes | 112 | * | * | * |
| 06 | Alpse-Maritimes | 773 | 0.33 | 0.26 | 44.9 |
| 07 | Ardèche | 313 | 0.4 | 0.25 | 41.2 |
| 08 | Ardennes | 376 | 0.5 | 0.31 | 38.1 |
| 09 | Ariège | 121 | * | * | * |
| 10 | Aube | 361 | 0.28 | 0.24 | 41.3 |
| 11 | Aude | 274 | 0.72 | 0.4 | 34.6 |
| 12 | Aveyron | 243 | 0.32 | 0.26 | 43.5 |
| 13 | Bouches-du-Rhône | 1795 | 0.56 | 0.34 | 42.7 |
| 14 | Calvados | 781 | 0.47 | 0.34 | 40.9 |
| 15 | Cantal | 164 | * | * | * |
| 16 | Charente | 374 | 0.56 | 0.3 | 38 |
| 17 | Charente-Maritime | 559 | 0.46 | 0.34 | 39.1 |
| 18 | Cher | 370 | 0.44 | 0.27 | 42.7 |
| 19 | Corrèze | 219 | 0.52 | 0.43 | 38.6 |
| 20 | Corse | 236 | 0.5 | 0.22 | 46.4 |
| 20 | Corse | 236 | 0.5 | 0.22 | 46.4 |
| 21 | Côte-d’Or | 549 | 0.42 | 0.33 | 42.2 |
| 22 | Côtes-d’Armor | 590 | 0.27 | 0.28 | 43.3 |
| 23 | Creuse | 102 | * | * | * |
| 24 | Dordogne | 337 | 0.25 | 0.24 | 39.4 |
| 25 | Doubs | 635 | 0.36 | 0.29 | 47.7 |
| 26 | Drôme | 435 | 0.42 | 0.29 | 39.5 |
| 27 | Eure | 738 | 0.4 | 0.26 | 42.6 |
| 28 | Eure-et-Loire | 506 | 0.48 | 0.32 | 42 |
| 29 | Finistère | 979 | 0.39 | 0.22 | 44.6 |
| 30 | Gard | 577 | 0.63 | 0.32 | 38.3 |
| 31 | Haute-Garonne | 949 | 0.49 | 0.31 | 43.1 |
| 32 | Gers | 136 | * | * | * |
| 33 | Gironde | 1304 | 0.42 | 0.28 | 41.6 |
| 34 | Hérault | 788 | 0.6 | 0.34 | 37.3 |
| 35 | Ille-et-Vilaine | 1036 | 0.36 | 0.3 | 42.5 |
| 36 | Indre | 235 | 0.67 | 0.38 | 37.1 |
| 37 | Indre-et-Loire | 597 | 0.63 | 0.38 | 40.5 |
| 38 | Isère | 1217 | 0.39 | 0.26 | 43.7 |
| 39 | Jura | 269 | 0.37 | 0.31 | 47.3 |
| 40 | Landes | 326 | 0.43 | 0.31 | 40.6 |
| 41 | Loir-et-Cher | 357 | 0.56 | 0.3 | 40.6 |
| 42 | Loire | 901 | 0.4 | 0.31 | 43.2 |

*Notes:* * Insufficient number of observations ($< 200$).

**Table F.9:** Department-Level Intergenerational Mobility Estimates *(continued)*

| | Department | Observations | IGE | RRC | AUM |
|---|---|---|---|---|---|
| 43 | Haute-Loire | 194 | * | * | * |
| 44 | Loire-Atlantique | 1467 | 0.44 | 0.27 | 41.3 |
| 45 | Loiret | 706 | 0.6 | 0.39 | 40.3 |
| 46 | Lot | 137 | * | * | * |
| 47 | Lot-et-Garonne | 319 | 0.7 | 0.29 | 40 |
| 48 | Lozère | 63 | * | * | * |
| 49 | Maine-et-Loire | 931 | 0.48 | 0.35 | 38.7 |
| 50 | Manche | 566 | 0.51 | 0.3 | 42.3 |
| 51 | Marne | 676 | 0.35 | 0.33 | 42.2 |
| 52 | Haute-Marne | 263 | 0.63 | 0.42 | 38.4 |
| 53 | Mayenne | 329 | 0.58 | 0.35 | 39 |
| 54 | Meurthe-et-Moselle | 862 | 0.57 | 0.33 | 41.5 |
| 55 | Meuse | 238 | 0.31 | 0.27 | 42.7 |
| 56 | Morbihan | 778 | 0.4 | 0.28 | 42.5 |
| 57 | Moselle | 1274 | 0.48 | 0.34 | 44 |
| 58 | Nièvre | 251 | 0.3 | 0.2 | 43.4 |
| 59 | Nord | 3668 | 0.61 | 0.39 | 37.4 |
| 60 | Oise | 1008 | 0.46 | 0.31 | 41.5 |
| 61 | Orne | 357 | 0.6 | 0.35 | 36.7 |
| 62 | Pas-de-Calais | 2145 | 0.71 | 0.44 | 34.4 |
| 63 | Puy-de-Dôme | 664 | 0.42 | 0.31 | 42.9 |
| 64 | Pyrénées-Atlantiques | 571 | 0.45 | 0.26 | 45.8 |
| 65 | Hautes-Pyrénées | 209 | 0.44 | 0.22 | 41.4 |
| 66 | Pyrénées-Orientales | 356 | 0.73 | 0.36 | 36.6 |
| 67 | Bas-Rhin | 1033 | 0.52 | 0.34 | 44.6 |
| 68 | Haut-Rhin | 792 | 0.53 | 0.32 | 46.8 |
| 69 | Rhône | 1583 | 0.45 | 0.28 | 44.7 |
| 70 | Haute-Saône | 273 | 0.83 | 0.34 | 39.4 |
| 71 | Saône-et-Loire | 661 | 0.61 | 0.38 | 42.1 |
| 72 | Sarthe | 635 | 0.53 | 0.37 | 38.9 |
| 73 | Savoie | 430 | 0.41 | 0.27 | 45.8 |
| 74 | Haute-Savoie | 629 | 0.4 | 0.21 | 54.7 |
| 75 | Paris | 1352 | 0.51 | 0.31 | 52.3 |
| 76 | Seine-Maritime | 1547 | 0.49 | 0.35 | 40.9 |
| 77 | Seine-et-Marne | 1529 | 0.41 | 0.27 | 45.7 |
| 78 | Yvelines | 1645 | 0.47 | 0.29 | 49.7 |
| 79 | Deux-Sèvres | 376 | 0.41 | 0.35 | 40.6 |
| 80 | Somme | 737 | 0.49 | 0.36 | 37.8 |
| 81 | Tarn | 354 | 0.48 | 0.38 | 36.1 |
| 82 | Tarn-et-Garonne | 202 | 0.52 | 0.21 | 42.3 |
| 83 | Var | 773 | 0.56 | 0.33 | 39 |
| 84 | Vaucluse | 468 | 0.48 | 0.3 | 41.9 |
| 85 | Vendée | 627 | 0.36 | 0.22 | 41.7 |
| 86 | Vienne | 464 | 0.49 | 0.35 | 40.1 |

*Notes:* * Insufficient number of observations (< 200).

**Table F.9:** Department-Level Intergenerational Mobility Estimates *(continued)*

| | Department | Observations | IGE | RRC | AUM |
|---|---|---|---|---|---|
| 87 | Haute-Vienne | 357 | 0.5 | 0.39 | 38.4 |
| 88 | Vosges | 504 | 0.46 | 0.3 | 39.8 |
| 89 | Yonne | 388 | 0.35 | 0.29 | 41.2 |
| 90 | Territoire de Belfort | 172 | * | * | * |
| 91 | Essonne | 1302 | 0.44 | 0.3 | 48.6 |
| 92 | Hauts-de-Seine | 1248 | 0.49 | 0.32 | 49.6 |
| 93 | Seine-Saint-Denis | 1495 | 0.44 | 0.24 | 47.8 |
| 94 | Val-de-Marne | 1188 | 0.41 | 0.23 | 52 |
| 95 | Val-d'Oise | 1366 | 0.48 | 0.29 | 47.4 |

*Notes:* * Insufficient number of observations ($< 200$).

| Child income definition | IGE-RRC | RRC-AUM | IGE-AUM |
|---|---|---|---|
| Household income | 0.65 | −0.58 | −0.46 |
| Individual income | 0.71 | −0.55 | −0.45 |
| Individual wage | 0.69 | −0.42 | −0.28 |

*Notes:* See Figure 7 for corresponding maps.

**Table F.10:** Correlation Between Department-Level Intergenerational Mobility Measures

| | *Dependent variable:* | | |
|---|---|---|---|
| | Intergenerational Elasticity | Rank-Rank Correlation | Absolute Upward Mobility |
| | (1) | (2) | (3) |
| Density | 0.051 | −0.136 | 0.498*** |
| | (0.195) | (0.198) | (0.143) |
| Unemployment rate | 0.253** | 0.062 | −0.277*** |
| | (0.113) | (0.115) | (0.083) |
| Gini | 0.111 | 0.287 | −0.345** |
| | (0.181) | (0.183) | (0.132) |
| % HS graduates | −0.177 | −0.274 | 0.451*** |
| | (0.201) | (0.203) | (0.147) |
| Cultural amenities | −0.083 | −0.249 | 0.313** |
| | (0.164) | (0.166) | (0.120) |
| Constant | −0.030 | −0.025 | 0.006 |
| | (0.109) | (0.110) | (0.080) |
| Observations | 85 | 85 | 85 |
| $R^2$ | 0.105 | 0.082 | 0.519 |

*Notes:* All variables are standardized such that the regression coefficient corresponds to the correlation. See Appendix Table D.1 for variable definitions and data sources. *p<0.1; **p<0.05; ***p<0.01.

**Table F.11:** Multivariate Correlation Between Intergenerational Mobility and Department Characteristics

| | *Dependent variable:* Child household income rank | | | | |
|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) |
| Parents income rank | 0.287*** | 0.287*** | 0.287*** | 0.184*** | 0.159*** |
| | (0.005) | (0.005) | (0.005) | (0.007) | (0.011) |
| Mover ($\hat{\gamma}$) | 4.306*** | 4.336*** | 4.691*** | 4.602*** | 4.486*** |
| | (0.462) | (0.462) | (0.465) | (0.465) | (0.465) |
| Parents income rank × Mover ($\hat{\delta}$) | −0.019** | −0.019** | −0.021*** | −0.030*** | −0.030*** |
| | (0.008) | (0.008) | (0.008) | (0.008) | (0.008) |
| Constant | 35.132*** | 34.709*** | 33.551*** | 27.720*** | 28.832*** |
| | (0.418) | (0.432) | (1.603) | (1.882) | (1.924) |
| Birth cohort | ✓ | ✓ | ✓ | ✓ | ✓ |
| Gender | | ✓ | ✓ | ✓ | ✓ |
| Department FE | | | ✓ | ✓ | ✓ |
| Parents' education | | | | ✓ | ✓ |
| Parents' 2-digit occupation | | | | | ✓ |
| $\mathbb{E}[\hat{\gamma} + \hat{\delta} p_{p,i}] = \hat{\gamma} + \hat{\delta} \times 50.5$ | 3.35 | 3.38 | 3.63 | 3.09 | 2.97 |
| $\mathbb{E}[\hat{\gamma} + \hat{\delta} p_p \| p_p = 25]$ | 3.83 | 3.86 | 4.17 | 3.85 | 3.74 |
| $\mathbb{E}[\hat{\gamma} + \hat{\delta} p_p \| p_p = 75]$ | 2.88 | 2.91 | 3.12 | 2.35 | 2.24 |
| Observations | 64,572 | 64,572 | 64,572 | 64,572 | 64,572 |
| Adjusted $R^2$ | 0.086 | 0.086 | 0.089 | 0.102 | 0.108 |

*Notes:* *p<0.1; **p<0.05; ***p<0.01.

**Table F.12:** Intergenerational & Geographic Mobility - Department Ranks

| | *Dependent variable:* Child household income rank | | | | |
|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) |
| Parents income rank | 0.311*** | 0.311*** | 0.303*** | 0.196*** | 0.156*** |
| | (0.005) | (0.005) | (0.005) | (0.007) | (0.013) |
| Destination department (ref.: stayers) | | | | | |
| Low-income | 0.535 | 0.562 | 0.683 | 0.433 | 0.233 |
| | (0.556) | (0.556) | (0.556) | (0.553) | (0.552) |
| Medium-income | 13.455*** | 13.488*** | 12.745*** | 13.000*** | 13.066*** |
| | (0.662) | (0.662) | (0.665) | (0.662) | (0.661) |
| High-income | 24.458*** | 24.470*** | 23.671*** | 23.891*** | 23.785*** |
| | (1.071) | (1.071) | (1.070) | (1.068) | (1.067) |
| Parents income rank × Low-income | −0.037*** | −0.038*** | −0.036*** | −0.039*** | −0.037*** |
| | (0.009) | (0.009) | (0.009) | (0.009) | (0.009) |
| Parents income rank × Medium-income inc | −0.051*** | −0.052*** | −0.047*** | −0.062*** | −0.065*** |
| | (0.011) | (0.011) | (0.011) | (0.011) | (0.011) |
| Parents income rank × High-income | −0.051*** | −0.051*** | −0.049*** | −0.073*** | −0.075*** |
| | (0.015) | (0.015) | (0.015) | (0.015) | (0.015) |
| Constant | 32.411*** | 32.012*** | 34.838*** | 29.355*** | 31.238*** |
| | (0.399) | (0.413) | (1.541) | (1.812) | (1.889) |
| Birth cohort | ✓ | ✓ | ✓ | ✓ | ✓ |
| Gender | | ✓ | ✓ | ✓ | ✓ |
| Department FE | | | ✓ | ✓ | ✓ |
| Parents' education | | | | ✓ | ✓ |
| Parents' 2-digit occupation | | | | | ✓ |
| Observations | 64,572 | 64,572 | 64,572 | 64,572 | 64,572 |
| Adjusted $R^2$ | 0.153 | 0.153 | 0.159 | 0.171 | 0.177 |

*Notes:* *p<0.1; **p<0.05; ***p<0.01.

**Table F.13:** Intergenerational Mobility and Income Level in the Destination Department