IARIW 2025

IARIW 2025

Thursday, October 2 & Friday, October 3

Trends in Intergenerational Mobility in Vietnam

Chinh Hoang-Duc
(University of Manchester)

Paper prepared for the IARIW–World Bank–UEB/VNU Conference on “Improving Well-being Measurement in Data-challenged Environments in Developing Countries for Better Evidence-based Policies” October 2-3, 2025

Parallel Session 4B Measuring Mobility
Time Slot: Friday, October 3, 8:00-9:30 AM

Trends in intergenerational mobility in Vietnam

Chinh Hoang-Duc

University of Manchester (UK)

12 September 2025

Abstract

This paper provides the first comprehensive analysis of trends in intergenerational mobility in Vietnam, a rapidly developing economy where the evolution of economic opportunity is not well understood. Using five waves of the Vietnam Household Living Standard Survey (VHLSS) from 2010 to 2018, I analyse cohorts born between 1978 and 1993 using a rolling cohort design and employ a Heckman selection model to correct for coresidency bias. The main findings are that while relative mobility in Vietnam has increased for more recent cohorts, absolute upward mobility has declined. This increase in relative mobility is primarily explained by a compression in the children's earnings distribution, which has led to decreasing cohort-level inequality, and by changes in the education-earnings relationship following educational expansion. The analysis also reveals significant heterogeneity across gender, urban-rural areas, and North-South regions, finding that a narrowing of the rural-urban earnings gap is a key driver of the national mobility trend. By providing the first trend-based analysis of intergenerational mobility in Vietnam, this study contributes a nuanced perspective on the relationship between mobility, education, and inequality, and demonstrates a robust framework for analysing mobility trends with repeated cross-sectional data in developing countries.

"Like father, like son." - *Proverb*

"If the children surpasses their parents, their home is blessed."

- *Vietnamese proverb, translated*

1 Introduction

Intergenerational mobility, the degree to which an individual's socioeconomic outcomes are independent of their parental background, is a fundamental measure of equality of opportunity. Often described as "the hope of economic development and the mantra of a good society" (Iversen et al., 2021), high mobility allows individuals to succeed based on merit rather than birthright, as exemplified by the post-war experiences of countries like the United States and the United Kingdom (Major and Machin, 2018). From an economic perspective, low mobility is not only a matter of social justice but also of efficiency, as it can lead to a misallocation of talent and perpetuate inequality, thus hindering long-term growth (Munoz and Van der Weide, 2025). The central captivating concept in the field is the "Great Gatsby Curve", which documents a negative relationship between cross-sectional inequality and intergenerational mobility (Corak, 2013; Durlauf et al., 2022). This connection implies that differences in mobility today can generate different inequality outcomes in the future and vice versa (Solon, 1999). This dynamic establishes why understanding intergenerational mobility is a primary concern for researchers and policymakers.

The extensive literature on intergenerational mobility, however, has historically been concentrated in developed countries, particularly in North America and Europe (Solon, 1999; Black and Devereux, 2011). In contrast, intergenerational mobility remains an under-researched area in developing nations, which are often characterised by rapid economic change and higher baseline inequality. While recent years have seen a growth in studies on developing regions such as Africa (Alesina et al., 2021) and China (Fan et al., 2021), alongside the development of new global databases that expand coverage to encompass nearly the entire world's population (Van der Weide et al., 2024; Munoz and Van der Weide, 2025), this progress is systematically hampered. Research in developing countries is constrained by severe data limitations that can bias estimates and complicate cross-country comparisons.

Researchers in developing economies face several data-related challenges. The most significant is the scarcity of long-term panel datasets that track parent-child pairs at comparable ages (Emran et al., 2018). This scarcity forces most studies to rely on "snapshot" estimates from cross-sectional or short-panel data, which are susceptible to several well-documented biases. First, coresidency bias arises when household surveys miss children who have moved away from the parental home, leading to a truncated sample of individuals who may be systematically different from those not observed (Emran and Shilpi, 2021). Second, lifecycle bias occurs when using single-year or short-run income snapshots to approximate lifetime earnings, which can severely attenuate estimates of intergenerational persistence due to the hump-shaped nature of age-earnings profiles Haider and Solon (2006). In response, the literature has developed methodological innovations, such as the use of rank-based measures that are more robust to lifecycle bias (Nybom and Stuhler, 2017) and bias-correction techniques like Heckman selection models to address sample selection issues (Fan et al., 2021), which this paper will adopt.

While snapshot estimates of intergenerational mobility are valuable, they are insufficient for understanding a country undergoing profound structural transformation. For an economy like Vietnam, which has experienced radical changes since the *Doi Moi* reforms of 1986, a static measure cannot reveal whether equality of opportunity is improving or deteriorating over time. Existing studies on Vietnam have provided valuable cross-sectional estimates of mobility (Nguyen and Nguyen, 2020; Dang, 2020), but these snapshots are often unreliable and contradictory. For example, studies examining mobility in Vietnam during the same period (circa 2010-2014) produced widely varying estimates due to different methodological choices (Doan and Nguyen, 2016; Dang, 2020; Nguyen and Nguyen, 2020), a phenomenon of inconsistent estimates also documented in other contexts (Iversen et al., 2021). Given these inconsistencies, this paper argues that a trend-based analysis offers a more robust and meaningful approach. Focusing on the direction of change across cohorts is more informative for policy and less sensitive to the methodological choices that plague single-point estimates. This leads to the central research gap this paper addresses: there is a notable absence of a comprehensive, multi-measure study on the trends of intergenerational mobility in Vietnam, particularly for cohorts born during the country's period of significant economic and educational expansion.

This paper fills this gap by providing the first detailed analysis of trends in intergenerational earnings and educational mobility in Vietnam for cohorts born between 1978 and 1993. Using five waves of the Vietnam Household Living Standard Survey (VHLSS) from 2010 to 2018, I adopt a rigorous empirical strategy designed to overcome the empirical challenges. The analysis uses a rolling birth cohort design to track trends over cohorts, applies a Heckman selection model to correct for coresidency bias, and controls for lifecycle effects with age polynomials. Furthermore, this study adopts a comprehensive set of mobility measures, including relative non-directional (IGE, Rank-Rank Slope, IGC), absolute directional (Absolute Upward Mobility), and relative directional (transition probabilities) metrics, to provide a holistic analysis of intergenerational mobility (Deutscher and Mazumder, 2023). The main findings are that while relative mobility in Vietnam has increased for more recent cohorts, however, absolute upward mobility has declined. This increase in relative mobility is primarily explained by a compression in the children's earnings distribution, which has led to decreasing cohort-level inequality, and changes in the education-earnings relationship following educational expansion. The analysis also reveals significant heterogeneity across gender, urban-rural areas, and North-South regions, finding that a narrowing of the rural-urban earnings gap is a key driver of the national mobility trend.

This study makes several contributions to the literature. Empirically, it moves the conversation on mobility in Vietnam beyond conflicting snapshot estimates to provide the first dynamic, longitudinal account of how mobility has evolved. It also offers a crucial case study of a country with stable aggregate inequality but dynamic mobility trends, challenging the field's dominant research narrative. Methodologically, it demonstrates a robust and replicable framework for analysing intergenerational mobility trends using repeated cross-sectional data, offering a practical solution to the data challenges in developing countries. Finally, by examining the inequality-mobility link in a dynamic, within-country context and decomposing trends into educational and inequality channels, this paper contributes to the Great Gatsby Curve literature and offers specific insights for policymakers aiming to foster inclusive growth.

The remainder of this paper is structured as follows. Section 2 details the data and discusses the key empirical challenges of coresidency and lifecycle bias. Section 3 introduces the various measures of intergenerational mobility used in the analysis. Section 4 presents the main empirical results, outlining the national trends and exploring heterogeneity across key subgroups. Section 5 provides a detailed discussion of the roles of education and inequality as the primary drivers of the observed trends. Section 6 concludes.

2 Data and Empirical challenges

2.1 Vietnam Household Living Standard Survey

Vietnam Household Living Standard Survey (VHLSS) is a biennially repeated cross-sectional survey of households in Vietnam conducted by the General Statistics Office of Vietnam (GSO) since 2002.[1] The survey is designed to be representative at the national, regional, urban-area, and provincial levels. The included households are sampled via a multi-step stratification. Since 2010, the master sample includes 3,133 communes, from each commune, three enumeration areas (EAs) are selected based on the number of households in the most recent population census. Then, households are randomly sampled from each EA and weighted accordingly. The VHLSS contains the basic household information such as size, housing, or poverty status, as well as detailed data on household income, consumption, and production activities. In this study, I focus on individual-level data including demographics (age, education, gender, marital status,...) and labour earnings.

This study uses data from five waves of VHLSS from 2010 to 2018 which originally included about 83,000 households. The choice of data in this specific window relates to the questionnaire design required to address coresidency bias, which will be discussed in a later section. Given the information about the relationship between household members, I identify the child-parent pairs. Since this paper focuses on labour earnings, and to be consistent with previous studies, I restrict the sample to only include children above 22 years old with parents under 65. This is because most people finish their education by age 22 and retire by 65. The

[1] This was previously the Vietnam Living Standard Survey (VLSS) conducted with the help of the World Bank in 1992-3 and 1997-8.

final dataset contains 22,845 child-parent pairs of children born between 1970 and 1995.[2] I retain the individuals who have full education, job, and work-from-afar[3] status. People with zero earnings are kept in the sample and their earnings are subsequently imputed using a Heckman selection model. Due to the small sample sizes in each birth cohort, and to smooth out cohort trends, in all the subsequent analyses, I use five-year rolling birth cohorts. This means, children born between 1976 and 1980 will be bundled into the 1978 rolling cohort, and children born between 1977 and 1981 will be bundled into the 1979 rolling cohort, and so on. This practice has been used in Mayer and Lopoo (2004). In addition to the national sample, this paper will also explore the differences in mobility trends between genders, areas (rural or urban), and regions (north or south). For robustness checks, analyses are also conducted on a subsample of children under 30 years of age and on a sample that excludes major metropolitan areas.

2.2 Empirical challenges

There are two major empirical challenges associated with this dataset.

Coresidency bias. Since VHLSS is a household survey from which I draw individual data, there is the possibility of coresidency bias, which means that children choosing to live with their parents (or vice versa) might be structurally different from those who do not. This has been pointed out as one of the main empirical problems of studying mobility in developing countries where longitudinal panel data are rare and surveys are often conducted at the household level (Van der Weide et al., 2024; Emran and Shilpi, 2021; Emran et al., 2018). Among many types of household divisions, for countries with a majority of the population living in the rural areas like Vietnam, there is a proportion of rural children moves to urban areas for work or studies. Although, it is very likely that these children will have better socio-economic outcomes, in absolute measures, than their parents, the implications

[2]Since the sample size for some older child cohorts is small, our main analyses only cover children born after 1975. That means the first rolling cohort is 1978. The effective sample size for the main analyses are 22,444. However, I still use all 22,845 observations to estimate the Heckman selection model.

[3]These are people not living in the household at time of survey. They are important in this study to address coresidency bias, which will be discussed later.

for intergenerational persistence are ambiguous. On the one hand, since these children are more likely to have different labour market decisions compared to their parents, intergenerational persistence for these families might be weaker. Truncating these households from the surveyed sample might lead the intergenerational persistence estimate to be higher than it actually is in the whole society. On the other hand, if there are intra-family factors that influence the decision to leave home which can also affect outcomes, such as abilities, ambitions, or parental resources, intergenerational persistence for these families can be higher. Then, not observing them would lead to attenuation bias. In a later section, I will discuss this bias in more detail with the data and propose a Heckman selection model to address the problem.

Lifecycle bias. Lifecycle bias arises when the mobility estimate is biased by the use of age-specific earnings rather than lifetime earnings. Haider and Solon (2006) documented that an age-earnings profile is hump-shaped; thus, estimates using snapshot earnings will vary as age varies. The VHLSS is not a panel dataset. Although some parts of the VHLSS were designed to be a panel, because the time span was short and the attrition rate was high (at least 50% by construction[4]), it is almost impossible to adopt the panel data methods used in, for example, Lee and Solon (2009); Mayer and Lopoo (2004) with the PSID data for the US, Blanden et al. (2007, 2004) with the BCS and NDCS data for the UK, or Fan et al. (2021) with the CFPS for China. Ideally, when estimating intergenerational earnings mobility, one would want to observe children's and parents' earnings throughout their lifetimes or, at least, during the same periods. The VHLSS, given its structure, does not allow for these observations. As the sample of children and parents is restricted to ages between 23 and 64, I can only observe child-parent pairs with children at their early career stages and parents at their late career stages. Estimations of earnings intergenerational persistence with this data will be subject to both lifecycle bias and attenuation bias due to measurement errors and transitory shocks to earnings (Black and Devereux, 2011; Solon, 1999). To address the former bias, given the structure of the data, I follow the standard practice and include

[4]The VHLSS was designed to have 50% of the sample being repeatedly surveyed during two separate periods: 2002-2008 and 2010–2018. However, since the GSO changed the master sample from which the VHLSS is drawn in 2009, one cannot construct a panel from 2002 to 2018.

quadratic terms of child and parent age in all of the persistence estimates. I also use a subset of children aged 30 and under to address the fact that earlier cohorts are older. Since the dataset is a repeated cross-section, I cannot explicitly address the latter bias. However, by using rolling birth cohorts, I can use information from nearby birth cohorts to smooth out the earnings measurements.

2.3 Variable definitions

Earnings. In the VHLSS, labour earnings are the only individual-level economic outcome. For each household, the VHLSS asks detailed questions about the jobs and wages of each individual, which cover the last three jobs at the time of the survey. These jobs can be wage jobs or self-employed jobs as long as they return recorded monetary benefits. The main question for wages is: "In the past 12 months, how much have you received in wages or salary, including the value of in-kind payments, from this job?". The VHLSS also asks questions about other job-specific benefits: "In the past 12 months, in addition to wages or salary from this job, how much have you received in cash and in-kind payments from the following sources: (a) Holidays and festivals; and (b) Other benefits (e.g., bonuses, uniforms, lunch allowance, travel allowance, sick leave allowance, occupational accident compensation, housing, property, etc.)". It should be noted that these benefits are not governmental transfers but are non-wage compensation from companies or organisations. The labour earnings of each individual are calculated as the total of wages and benefits from up to three jobs in the last 12 months at the time of the survey. Since I pooled data from different waves, all earnings are adjusted to the 2022 price index. Earnings are only observed for individuals reporting to have jobs and to live in the surveyed households. This creates the selection bias that I discuss below.

Education. Education is captured via four variables: degrees, degree levels, vocational training, and years of school. The "degree" variable represents the highest degree a person has obtained, divided into an eight-level ordered scale: none, primary, secondary, high school, college diploma, bachelor's, master's, and doctorate. From this variable, I create three degree levels: low (no or primary education), medium (secondary and high school), and high (higher education, i.e., any formal education above high school). Vocational training is a dummy variable indicating if the surveyed individual holds a vocational certificate of any level. Lastly, "years of school" is the combination of finished grades (formal K-12 education) and obtained degrees. Based on average training times, I assign secondary vocational training an additional two years of school, vocational college an additional three years, academic college and university training four years, a master's degree two years, and a doctorate three years. Typically, for people to obtain a higher level of education, they have to finish the previous lower levels. Therefore, the final "years of school" variable ranges from 0 for individuals who never went to school to 21 for those who finished their doctoral training. Since years of K-12 education are only observed for individuals living in the household, I mainly use the degree variables in the analyses.

Work-from-afar. In addition to information on members currently living in the household, the VHLSS also collected some information about people who had left. These are people "who have migrated, including export workers, for more than 6 months, to earn a livelihood" for the household, or "former members who (1) left the household within the last 10 years or (2) may have left at an earlier time but the household still considers them important, in terms of obligations to old parents living with your household or financial support to your household". In this paper, I refer to these people as work-from-afar (WFA). The VHLSS collected information from WFA individuals by asking another household member, usually the household head. Questions asked include gender, relationship with the household head, year of birth, marital status, highest degree, and vocational training.

Table 1 presents descriptive statistics of the main variables used in this paper, disaggregated by children's rolling birth cohorts. On average, children are around 30 years old while their parents are around 55 at the time of the survey; by construction, both children and parents in later cohorts are younger. While children in later cohorts tend to have more education, this trend is not clear among their parents. For children, the share of individuals working fluctuates around 90%, while for parents, this figure increases from 36% for the 1978 child cohort to 74% for the 1993 child cohort. Observed earnings of children decrease across cohorts while the earnings of parents increase. Combined with the ages of children and parents in each cohort, these trends indicate a hump-shaped age-earnings profile, where earnings at earlier and later ages are lower than those at middle ages.

Table 1: Descriptive statistics by children cohort

	1978		1983		1988		1993	
Statistic	Avg	Std. Dev.	Avg	Std. Dev.	Avg	Std. Dev.	Avg	Std. Dev.
Children								
Age	34.962	(3.168)	30.767	(3.322)	27.300	(2.610)	24.615	(1.340)
Female (%)	38.471	(48.667)	40.237	(49.042)	42.325	(49.411)	41.283	(49.238)
Years of School	10.862	(4.677)	11.629	(4.216)	11.779	(3.997)	11.704	(3.803)
Education: Low (%)	32.283	(46.769)	22.987	(42.079)	19.238	(39.419)	16.814	(37.401)
Education: Medium (%)	38.120	(48.582)	47.063	(49.919)	50.138	(50.003)	53.248	(49.898)
Education: High (%)	29.597	(45.661)	29.949	(45.808)	30.625	(46.096)	29.938	(45.802)
Vocational (%)	14.594	(35.315)	20.206	(40.158)	16.387	(37.019)	13.638	(34.322)
Has Job (%)	92.936	(25.629)	93.504	(24.647)	90.112	(29.851)	87.824	(32.702)
WFA (%)	26.095	(43.928)	22.190	(41.556)	20.163	(40.124)	22.335	(41.652)
Earnings (mil VND)	58286.015	(74979.855)	56290.195	(62576.136)	50072.282	(51004.159)	45018.501	(45967.997)
Log Earnings	11.090	(0.820)	11.019	(0.799)	10.953	(0.761)	10.932	(0.742)
Predicted Earnings (mil VND)	75203.305	(31596.241)	72310.633	(30970.354)	65400.685	(25474.354)	60246.072	(18414.197)
Log Predicted Earnings	11.144	(0.409)	11.103	(0.414)	11.016	(0.380)	10.962	(0.293)
Parents								
Age	59.145	(3.659)	56.997	(4.541)	54.437	(5.083)	52.089	(5.338)
Female (%)	60.887	(48.815)	52.625	(49.936)	47.875	(49.958)	46.080	(49.849)
Years of School	8.074	(4.556)	8.732	(4.412)	8.668	(4.363)	8.447	(4.389)
Education: Low (%)	48.920	(50.003)	41.988	(49.359)	40.613	(49.114)	42.720	(49.470)
Education: Medium (%)	40.689	(49.140)	47.161	(49.924)	48.612	(49.984)	46.897	(49.907)
Education: High (%)	10.391	(30.523)	10.852	(31.106)	10.775	(31.008)	10.383	(30.507)
Vocational (%)	11.967	(32.467)	15.578	(36.268)	12.275	(32.817)	9.198	(28.901)
Has Job (%)	36.077	(48.036)	48.172	(49.971)	64.125	(47.966)	74.358	(43.669)
WFA (%)	1.343	(11.513)	1.011	(10.006)	1.163	(10.720)	0.988	(9.893)
Earnings (mil VND)	14741.238	(42708.754)	21863.162	(43270.027)	33516.854	(47210.711)	44854.557	(52043.971)
Log Earnings	10.087	(1.135)	10.226	(1.125)	10.424	(1.095)	10.642	(0.995)
Predicted Earnings (mil VND)	28091.005	(20843.257)	33647.302	(23687.741)	40271.964	(29762.590)	48140.945	(35899.848)
Log Predicted Earnings	10.035	(0.623)	10.226	(0.618)	10.401	(0.617)	10.585	(0.600)

Notes: Each year of birth is a 5-year *rolling* birth cohort centering the displayed year. That means, the 1978 cohort includes children born from 1976 to 1980 and so on. Parents are summarised by the rolling birth cohort of their children. The dataset includes 22,444 child-parent pairs. For the observed log earnings, there are 15,378 positive observations for children and 13,625 for parents. "WFA" stands for "working-from-away". Earnings are annual at 2022 prices.

2.4 Addressing coresidency bias

By construction of the VHLSS, since earnings are only observed for individuals who stay at the surveyed household and choose to work, these truncations in the earnings variable can reflect selection bias. As discussed in the preceding section on coresidency bias, these individuals might be structurally different from others, which means that dropping them from the sample would affect the intergenerational mobility estimates. While the same arguments apply to those who decide to work versus those who do not, at the end of this section, I will show that this employment decision does not constitute sample selection bias in the earnings data. Therefore, this section is devoted to addressing coresidency bias.

This problem fits into a classic Heckman sample selection model, in which the missing values are not at random but result from a latent choice process (Heckman, 1979). This model can be presented via a system of two structural equations. The first equation dictates individuals' choices to work from afar (WFA) or to stay home:

$$C_i^* = \gamma_0 + z_i\gamma_z + X_i\gamma_X + u_i, \quad C_i = \mathbb{1}(C_i^* > 0), \tag{1}$$

with C_i being a dummy variable that equals one when the individual chooses to stay at home (in reverse of WFA); C_i^* is its latent structural variable; z_i is the excluded variable that aids identification; X_i is a set of demographic, socio-economic, and geographic variables.

The second equation controls the outcome:

$$y_i^* = \beta_0 + X_i\beta_X + \varepsilon_i, \quad y_i = y_i^* \text{ only if } C_i = 1, \tag{2}$$

where y_i being the observed log of labour earnings.

The problem arises when the error terms (u_i, ε_i) are correlated if there are unobservables that affect both decision to WFA and outcome, such as ability or ambition, or more relevant to this study, parental resources. Then, estimating coefficients in Equation 2 using only observed units would lead to specification errors (Heckman, 1979). To address this problem, we follow the two-stage estimation procedure with the assumption that the error terms follow a bivariate normal distribution. In the first stage, we estimate the probit model using the full sample of individuals at home or WFA:

$$Pr(C_i = 1|z_i, X_i) = \Phi(\gamma_0 + z_i\gamma_z + X_i\gamma_X). \tag{3}$$

Then from the estimated coefficients $(\hat{\gamma}_0, \hat{\gamma}_z, \hat{\gamma}_X)$, I calculate the Inverse Mills ratio (IMR):

$$\hat{\lambda}_i = \frac{\phi(\hat{\gamma}_0 + z_i\hat{\gamma}_z + X_i\hat{\gamma}_X)}{\Phi(\hat{\gamma}_0 + z_i\hat{\gamma}_z + X_i\hat{\gamma}_X)}, \tag{4}$$

where ϕ and Φ are standard normal PDF and CDF. In the second stage, I correct for selection bias by including the calculated IMR in our outcome model. Specifically, I estimate the following model with OLS:

$$y_i = \beta_0 + X_i\beta_X + \hat{\lambda}_i\beta_\lambda + \varepsilon_i. \tag{5}$$

If there is selection bias, the coefficient corresponding to the IMR will be significantly different from zero. In that case, including the IMR in the outcome equation helps control for the unobservables that influence both WFA decision and earnings. Finally, with the corrected outcome coefficients $(\hat{\beta}_0, \hat{\beta}_X, \hat{\beta}_\lambda)$, I impute earnings for the full sample regardless of their decision to WFA or stay at home.[5] I repeat this procedure for children and their parents separately. The imputed earnings will be used to study intergenerational persistence. Figure A1 shows the distributions of observed and predicted earnings for both parents and children. While the means are roughly the same, the predicted earnings has more compressed distributions. This is a natural consequence of the model specifications, which strip away the unobserved heterogeneity between individuals. As I use these specifications consistently for all children and parents across cohorts, this should not affect the overall trend of the mobility estimates. Figure A2 also shows that the joint distribution between children and their parents using predicted earnings is also more compressed, though the general positive correlation remains.

In summary, the entire estimation process involves two major steps. Step one is devoted to correcting and imputing the missing earnings due to selection, which itself includes a

[5] I calculated the IMR for the unobserved units as: $\lambda_i = \frac{-\phi(\hat{\gamma}_0 + z_i\hat{\gamma}_z + X_i\hat{\gamma}_X)}{1-\Phi(\hat{\gamma}_0 + z_i\hat{\gamma}_z + X_i\hat{\gamma}_X)}$.

two-stage Heckman model estimation. Step two uses the imputed data to estimate the coresidency-bias-corrected measures of intergenerational mobility. Therefore, I bootstrap the entire two-step process to obtain correct standard errors for the final intergenerational estimates.

There are three points to note about the specification of these models. First, I choose the excluded variable (z_i) to be the share of households who have WFA members (other than the considered one) in the same commune. A higher communal WFA rate strongly associates with a higher probability of an individual leaving home for WFA (see the first two columns in Table A3). I chose the commune level due to the specific socioeconomic and geographical characteristics at this level. Households in the same commune often know about each other's economic activities and tend to adopt similar practices. This can be due to similar geographical and socioeconomic characteristics. Proximity allows for word-of-mouth communication and the build-up of social norms. Individuals in the same commune also benefit from similar economic development and local policies, which leads to the WFA decision being highly heterogeneous among communes. In the dataset, there are nearly 2,700 communes, with the number of households ranging from 2 to 44 and the WFA rate ranging from 0% to 95%. This variation helps weaken collinearity between the IMR and other X_i variables in the outcome model, which is required for identifying the structural parameters.

Second, the choice of the set X_i needs to be comprehensive. The variables I include are degree, vocational education, gender, age, survey wave dummies, area dummies (urban vs. rural), and region dummies (north vs. south). However, since the area and region variables are at the household level, using them to estimate, and later predict, earnings would lead to spurious correlations between children and parents, potentially overestimating intergenerational persistence. Therefore, I interact these geographical variables with all other individual characteristics to avoid using a common source of variation for both children and parents. Survey wave dummies are also at the household-level; however, as the model includes individual-specific age and the data are cross-sectional, these dummies act as a cohort effect, not a time effect. Therefore, survey waves are not a common source of variation between children and parents. With these covariates, my earnings model achieves an adjusted R^2 of

22.7% for children and 29.1% for parents (see the first two columns in Table A4).

Third, since earnings are only observed for individuals choosing to stay home and work, the two-stage procedure described above was effectively estimated using the subsample of working children and parents. This explains why the number of observations in Table A3 is not 22,845. However, in this dataset, missing data due to not working does not constitute selection bias. I show this by estimating a two-stage selection model similar to the one described above, using "having a job" as the selection variable and the ratio of small children in the family as the excluded variable.[6] The last two columns in Table A3 show that the small-children ratio is significantly associated with the decision to work for both (adult) children and their parents. However, the last two columns in Table A4 show no significant coefficient corresponding to the IMR of selection into working. This means that for the current model specification, missing data due to not working do not appear to be a source of selection bias. Therefore, I can use the estimates from my previous Heckman-corrected model for WFA and impute the missing earnings data for all individuals in my sample.

This Heckman-corrected procedure was also adopted by Fan et al. (2021) using a panel household survey in China; however, there are two main differences. The first concerns the outcome. While I use individual labour earnings, they use income, which is a combination of individual wages, transfers, and family-related income (e.g., rent from family properties). I avoid analysing income since, in the VHLSS, income, or many of its components, is a family-level variable. Given this, by construction, the income of parents and children would always be correlated, regardless of individual-specific traits like education or labour market earnings. Second, while I bootstrap the entire two-step process, Fan et al. (2021) only bootstrapped the second step of their intergenerational mobility estimation procedure. This leads to their main results' standard errors being too small compared to the estimates from

[6]This is different from the "children" I have referred to in this paper hitherto. Small children are those under age 16, dependent on their parents, while adult children are those above 22 and are the children in my child-parent pairs. For young adults of working age, having more children in the household is associated with a higher probability of working. This might be due to higher intra-household demand or higher motivation to earn. For older people, having more small children in the family is associated negatively with their probability of working. This might be due to the fact that the elderly in Vietnam are normally the childminders in their families. As younger labourers in the household have a higher comparative advantage in labour market production than the elderly, and the elderly have a higher comparative advantage in intra-household production than market childminders, this division of labour is optimal.

observed, uncorrected units, overestimating the significance of their results. By correctly bootstrapping the entire process, my intergenerational estimates from imputed earnings have roughly the same standard errors as the naive estimates from observed units (Figure A5), which allows this paper to interpret trends with a more statistically robust manner.

3 Measures of Intergenerational Mobility

This section introduces the different measures of intergenerational mobility used in this paper. Since mobility is a multidimensional issue and its measurement depends on one's normative objectives, analysing multiple measures is useful and can provide a more holistic understanding (Chetty et al., 2014a; Fields and Ok, 1999). As pointed out by Deutscher and Mazumder (2023) and Genicot et al. (2024), among many others, each measure of mobility bears a different meaning and can be highly contextual, which makes using only one measure inadequate. For example, a society in which many children have equal or lower earnings than their parents can be characterised as having very low child-parent persistence (meaning high relative mobility) but low absolute upward mobility. Alternatively, a society can have a different level of persistence for those at the left tail versus the right tail of the distribution. Moreover, in the context of lacking high-quality data, any single mobility measurement tends to be subject to certain biases. Having multiple measurements can reduce the dependence on one bias-prone result.

This paper reports many different intergenerational mobility statistics, most of which are on earnings and some on education. Following Genicot et al. (2024), I categorise these measures into four groups based on whether the measure is directional or non-directional and absolute or relative. Directional mobility measures "emphasize the direction of movement along ordered or ranked categories." This means the directions of movement have economic meaning and can be valued differently. Non-directional mobility, on the other hand, only measures movement or deviation and "places no weight at all on the category in question." Absolute mobility measures the changes that "stem from individual changes in intrinsic economic standing, independently of changes in the positions of others," while relative mobility is "only concerned with changes in comparative individual standings" (Genicot et al., 2024). Table 2 classifies all analysed statistics into their respective categories.

Table 2: Classification of Intergenerational Mobility Measures

	Non-directional	**Directional**
Relative	IGE, Rank-rank, Correlation, Education elasticity	Transition matrices, Transitional probabilities
Absolute	-	Absolute upward mobility

Intergenerational Elasticity.

Earnings. The intergenerational elasticity of earnings (IGE) is the most common estimate of intergenerational mobility. In its original form, the IGE is simply the coefficient resulting from regressing the log of children's earnings on the log of their parents'. In this paper, to mitigate the lifecycle bias arising from using snapshot earnings, I control for quadratic terms of children's and parents' ages. This practice is common in the literature (Lee and Solon, 2009; Deutscher and Mazumder, 2023). The model to be estimated is:

$$y_i^c = \alpha_0 + \beta y_i^p + \alpha_1 age_i^c + \alpha_2 (age_i^c)^2 + \alpha_3 age_i^p + \alpha_4 (age_i^p)^2 + \epsilon_i,$$

or

$$\tilde{y}_i^c = \beta \tilde{y}_i^p + \varepsilon_i, \tag{6}$$

$$\text{where} \quad \begin{cases} \tilde{y}_i^c = y_i^c - [\hat{\delta}_1 age_i^c + \hat{\delta}_2 (age_i^c)^2 + \hat{\delta}_3 age_i^p + \hat{\delta}_4 (age_i^p)^2] \\ \tilde{y}_i^p = y_i^p - [\hat{\gamma}_1 age_i^c + \hat{\gamma}_2 (age_i^c)^2 + \hat{\gamma}_3 age_i^p + \hat{\gamma}_4 (age_i^p)^2] \end{cases}$$

where y_i represents the log earnings of children $(.^c)$ and parents $(.^p)$ and $\tilde{y}_i$ is the residual from regressing on age and its quadratic terms. This equation is estimated for each five-year rolling child birth cohort.[7] A higher IGE value indicates less mobility and a greater

[7] All the measures introduced in this section are estimated for each five-year rolling child birth cohort, unless otherwise stated.

dependence of children's outcomes on their parents'. For instance, an IGE of 0.3 means that a 10 percent point difference in income between two families in the parent generation is expected to result in a 3 percent point difference in the children's generation. In the taxonomy of Genicot et al. (2024), the IGE is a relative measure because, within a sample, the IGE of one individual changes due not only to that child or his/her parent but also to any other pair.

Education. In developing countries, where earnings data are lacking, education is often the preferred outcome for intergenerational studies and is claimed to be less affected by measurement errors or lifecycle bias (Van der Weide et al., 2024; Alesina et al., 2021). The intergenerational elasticity of education (IGEdu) is measured as the coefficient from the regression of a child's education on their parent's:

$$e_i^c = \alpha^e + \beta^e e_i^p + \varepsilon_i^e, \tag{7}$$

where e_i is either degree level or years of schooling. Similar to the IGE, this measure captures the dependence of a child's education on their parents' education in levels.

Intergenerational Log Correlation. The IGE, when measured using snapshot earnings data, is subject to lifecycle bias, which can be reflected in the differences in earnings variance between children and parents or across different ages when earnings are captured (Blanden et al., 2004; Black and Devereux, 2011). The literature often addresses this problem by using the intergenerational correlation (IGC) alongside the IGE, which is the Pearson correlation between children's log earnings and their parents'. The IGC shows the simple strength of the relationship between children and their parents. In this paper, to be consistent with the IGE measurement described above and to correct for the portion of earnings that can be explained by age, I use the partial correlation between child and parental earnings, controlling for age. The IGC is:

$$r = \frac{cov(\tilde{y}^c, \tilde{y}^p)}{\sigma_{\tilde{y}^c}\sigma_{\tilde{y}^p}} = \beta\frac{\sigma_{\tilde{y}^p}}{\sigma_{\tilde{y}^c}} \tag{8}$$

where $\sigma_{\tilde{y}}$ is the sample standard deviation of the residual from regressing earnings on age, which represents inequality. As inequality can change from one cohort to another, Blanden et al. (2004) termed the IGC the "coefficient adjusted for inequality changes." In addition to adjusting for inequality changes, Emran et al. (2018) found that the IGC can mitigate the bias in the IGE caused by using coresident data.

Intergenerational Rank Correlation. The intergenerational rank correlation, or the rank-rank (R-R) slope estimate, has gained popularity since Chetty et al. (2014a). Although a relative, non-directional measure, the R-R focuses on the relative earnings position of children and parents compared to their peers rather than on earnings magnitude. Thus, it factors out the effect of economic expansion and focuses only on the ability of each generation to climb the "social ladder." It is a measure of positional mobility that is not affected by changes in inequality. An R-R value of 0.5 means that if two parents are 20 percentiles apart in their earnings ranks, their children will be 10 percentiles apart. I estimate the R-R by regressing the percentile rank of children's earnings, after partialling out quadratic age controls, on that of their parents to mitigate lifecycle bias. This approach to a quantile-related mobility measure is similar to that of Charles and Hurst (2003). The R-R is the θ coefficient in the following equation:

$$rank(\tilde{y}_i^c) = \theta rank(\tilde{y}_i^p) + u_i. \tag{9}$$

Similar to other measures, the R-R is estimated for each child rolling cohort. There are two points that need to be made clear about this. First, both child rank and parent rank are calculated among children and parents within the same child cohort. This means children are compared to other children born around the same time, and parents are compared to other parents having kids around the same time. Second, percentile ranks are calculated for each data subset. This is applied to both rolling cohort subsets and, later, for subgroup analyses. An R-R estimated from subset ranks has a different meaning than an R-R estimated from full dataset ranks. While the former retains the pure meaning of relative, positional mobility within a subgroup, the latter is affected by the absolute position of units in that subgroup compared to the full dataset (Deutscher and Mazumder, 2023). Since this paper focuses on the changes in the pure relationship between children and parents, I choose to estimate the R-R using subgroup ranks. In addition to providing another perspective on intergenerational persistence to the IGE and IGC, the R-R is also proven to be less prone to measurement error, transitory income fluctuations, and lifecycle bias (Nybom and Stuhler, 2017).

Absolute Upward Mobility. This absolute, directional measure is perhaps the most intuitive measure of IM to the public and became highly influential after the publication of Chetty et al. (2017)'s paper. Absolute upward mobility (AUM) calculates the share of children who earn more than their parents:

$$AUM = \frac{1}{n} \sum_{i} \mathbb{1}(y_i^c > y_i^p). \tag{10}$$

Although a clear and intuitive measure, Genicot et al. (2024) point out that AUM does not account for the magnitude of a child's income gain over their parent's, treating a minor increase identically to a substantial one. AUM is also sensitive to measurement errors and lifecycle bias. Due to the data structure, this paper cannot directly address these problems. Later, I will calculate the AUM and other statistics using data on children aged 30 and under to partially mitigate the lifecycle bias stemming from the fact that children of earlier cohorts are older than those of recent cohorts.

Transitional matrices.

Earnings. For each five-year child rolling cohort, I construct a transition matrix that shows the percentage of children in each earnings quintile (Q1, Q2, Q3, Q4, Q5) conditional on their parents' earnings quintiles, all after partialling out the age quadratic terms. Q1 represents the lowest 20 percent earnings group, and Q5 indicates the top 20 percent of earners. Similar to when calculating percentile ranks, children are compared to each other in the same cohort, and parents are compared to other parents who gave birth in that same cohort. A transition matrix is a relative, directional measure that shows the likelihood of children achieving certain relative earnings groups given their parents' relative positions. In theory, a fully mobile society would have all the cells of this matrix equal to 20 percent. Depending on the cell location, its value or its difference from 20 can have a different meaning. For example, if 40 percent of children of Q1 parents stay at Q1, this can be a negative signal of poverty persistence, but if 40 percent of children of Q2 parents reach Q3, this can be a positive signal of upward mobility. These concepts will be explored further in the discussion of transitional probabilities.

Education. The same general idea can be applied to calculate a transition matrix for education. Instead of breaking individuals into different quantiles of education, I use the three degree-level variables introduced above. The cells of the matrix are the percentage of children in each degree level given the level of their parents. I illustrate these values using a mosaic plot similar to that in Alesina et al. (2021).

Transitional probabilities. The transition matrix allows us to understand relative, directional mobility by examining each of its cells. Since this paper focuses on trends, I choose to report eight different cell statistics over cohorts. Similar to the R-R and transition matrices discussed above, all "earnings" mentioned in the mobility statistics related to the quantiles below are residuals after partialling out the quadratic age terms.

Persistence at Top ($Q_5 \rightarrow Q_5$) and Immobility at Bottom ($Q_1 \rightarrow Q_1$). These statistics examine the two tails of the distribution more closely. "Persistence at Top" is the share of children who are at Q5 given that their parents are at Q5. "Immobility at Bottom" is the share of children who stay at Q1, like their parents. These two measures show the level of intergenerational persistence among two different groups of people.

Upward Mobility ($Q_x \rightarrow Q_{x+1}$) and Downward Mobility ($Q_x \rightarrow Q_{x-1}$). These statistics show the magnitude of directional mobility. The former is the share of children whose earnings are one quintile higher than their parents', and the latter is the share who earn one quintile less. On the transition matrix, "Upward Mobility" is represented by the cells directly to the right of the diagonal, and "Downward Mobility" is represented by the cells on the left. These reflect the degree of short-range movement in society.

Upward from Middle ($Q_3 \rightarrow Q_{>3}$) and Downward from Middle ($Q_3 \rightarrow Q_{<3}$). These statistics focus on individuals with an average background. "Upward from Middle" is the share of children whose earnings are in the 4th or 5th quintiles given that their parents are at the 3rd quintile. These can be considered individuals with a middle-class origin who strive for the upper classes. "Downward from Middle", on the other hand, reports the share of children from median families who fall to a lower earnings rank in society.

Extreme Upward Mobility ($Q_1 \rightarrow Q_5$) and Extreme Downward Mobility ($Q_5 \rightarrow Q_1$). These two statistics reflect the degree of long-range movement. "Extreme Upward Mobility" is the share of children leaping to the Q5 earnings group when their parents are at Q1, and "Extreme Downward Mobility" is the share of individuals falling to the bottom group from top-group families.

4 Trends in Intergenerational Mobility

This section reports and discusses the trends in intergenerational mobility in Vietnam via the different measures introduced above. I examine the trends by birth cohort, specifically, a five-year child birth cohort, similar to the approach in Davis and Mazumder (2024); Fan et al. (2021); Chetty et al. (2017); Mayer and Lopoo (2004), among many others. This approach differs from that of papers studying trends by calendar year, such as Lee and Solon (2009); Nguyen and Nguyen (2020); Brandén et al. (2024). While there is no right or wrong approach, the former is more suitable for my short, repeated cross-sectional dataset and my goal of understanding how recent cohorts fare compared with previous ones. For each rolling cohort data subset, I calculate the aforementioned mobility statistics. This differs from the approach of Lee and Solon (2009), who estimate the IGE with time dummy interactions using a pooled sample. Since I examine many different mobility measures, estimating them for each cohort separately ensures that the estimations are consistent with each other.

4.1 National trends

Figure 1 shows the IM trends in Vietnam for rolling cohorts between 1978 and 1993[8] with four measures: IGE, R-R, IGC, and AUM. In general, relative, non-directional intergenerational persistence tends to decrease over birth cohorts, albeit with different magnitudes. The IGE shows the most significant and continuous decline, from 0.425 for the 1978 rolling cohort to 0.289 for the 1993 cohort (see Table A1). This means that recent cohorts' earnings are less dependent on their parents' than those of prior cohorts, which equates to higher IM. The R-R and IGC also experience decreases, but with a much smaller magnitude, regardless of being in the same category of measure as the IGE. The R-R measure ranges from 0.606 for the 1993 cohort to 0.664 for the 1982 cohort. Although the R-R point estimates follow a slight decrease over cohorts, this trend is only weakly statistically significant. Specifically, the biggest change in the R-R is between 1982 and 1993, with a decrease of 0.058 ($SE \approx 0.028; p \approx 0.041$).[9] This means the rank dependence of children on their parents also decreases slightly for recent cohorts. Similarly, the IGC also experiences gradual decreases, from a high of 0.669 for the 1979 and 1982 cohorts to a low of 0.603 for the 1993 cohort (difference is 0.066, $p \approx 0.004$).

It is immediately noticeable that the IGE is smaller than the other two measures for all cohorts. This is because the R-R and IGC are not as sensitive to biases that attenuate the estimated values as the IGE is (Nybom and Stuhler, 2017; Emran et al., 2018). Moreover, the differences between the IGE and the R-R or IGC in both values and trends have implications for changes in inequality across cohorts. A decreasing IGE means that the transmission of the gap in parental earnings to the gap in children's earnings has become weaker, implying a faster regression to the mean. However, a mostly flat R-R (slight decrease) implies that the magnitude of dependence of children's ranks on their parents' ranks remains roughly the same. Given that, it is possible that children's earnings are becoming more equal and

[8]The 1978 rolling cohort includes those born between 1976 and 1980. The 1993 rolling cohort includes those born between 1991 and 1995.

[9]I compute the standard error of the difference between the two independent estimates as $\sqrt{SE_1^2 + SE_2^2}$ and approximate the p-value using $df = 1000$.

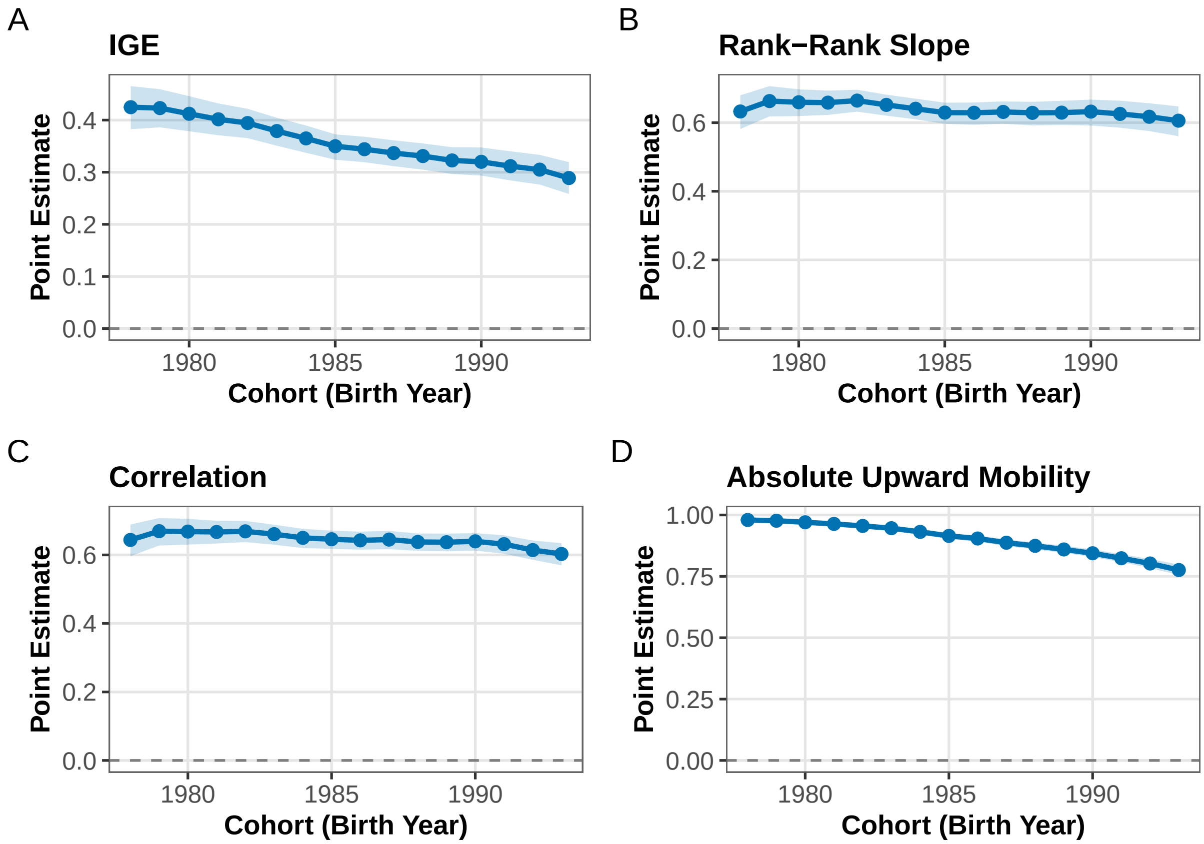

Figure 1: Trends in Intergenerational Earnings Mobility

Notes: Earnings used in these panels are corrected via a Heckman-selection model. Shaded areas are 95% bootstrap confidence intervals with 1000 iterations.

compressed, but these changes in levels have not significantly affected the relative positions of children; thus, the R-R does not change as much.

The IGC trend confirms this intuition. As the IGC scales the IGE up by the ratio of inequality between parents and children (Equation 8), changes in inequality among children and parents across cohorts will differentiate the IGE and IGC trends. Figure 7 plots the trends in the Gini coefficient of earnings separately for parents and children. While inequality among parents remains stable, inequality among children has reduced across cohorts. These trends hold even after adjusting for age in both sets of earnings.

While this compression signals a more equal earnings distribution among children, it is not necessarily the case that recent cohorts enjoy better shared prosperity. The last panel in

Figure 1 plots the AUM, showing that fewer children have earnings surpassing their parents' in recent cohorts. Summary statistics from Table 1 also point out that while parental earnings increase gradually over cohorts, children's earnings stagnate or even decline. However, this could be due to lifecycle effects: earnings at earlier or later years tend to be smaller than earnings in mid-life years, forming a hump-shaped age-earnings profile (Haider and Solon, 2006). Children in recent cohorts are only in their 20s, at the beginning of the hump, whereas parents in recent cohorts are closer to the peak compared to parents in earlier cohorts. To adjust for this, I restrict the sample to only children aged 30 and under to calculate intergenerational mobility. This allows for a comparison of children in different cohorts while keeping their ages roughly similar. Figure A9 shows that the trends of all main measures remain: relative non-directional mobility increases, but AUM decreases. While this practice cannot fully address lifecycle bias, restricting the sample exposes all cohorts to the same bias, making the trend analysis more credible.

Table 3: Intergenerational Earnings Transition Matrices

	Child's Quintile				
	Q1	Q2	Q3	Q4	Q5
Panel A: 1980					
Q1	54.72	22.55	17.13	4.72	0.87
Q2	27.50	34.85	19.79	15.06	2.80
Q3	15.06	22.42	26.80	21.54	14.19
Q4	2.45	15.94	24.17	30.30	27.15
Q5	0.35	4.20	12.08	28.37	54.99
Panel B: 1990					
Q1	48.53	26.67	13.39	9.91	1.50
Q2	31.49	29.09	20.79	13.82	4.81
Q3	15.44	25.36	25.06	21.81	12.32
Q4	4.03	15.02	25.36	29.21	26.38
Q5	0.54	3.85	15.38	25.24	54.99

Notes: Earnings used in this table are corrected via a Heckman-selection model. Each panel is a *rolling* cohort centering the displayed year. The rows are parental earnings quintile and the columns are child's quintiles.

The measures discussed above are what Deutscher and Mazumder (2023) term "global" measures, providing a general understanding of the child-parent economic relationship for all of society. I am also interested in "local" measures, which focus on different groups in the earnings distribution. Table 3 reports the transition matrices for the 1980 and 1990 rolling cohorts. The values on the main diagonals are much higher than the ideal 20% level, confirming the high persistence values of the R-R.[10] Specifically, for the 1980 cohort, 54.72% of children born to the lowest-earning parents are also in the lowest earning group themselves. On the other end, 55% of parents in the highest earnings quintile have children whose earnings fall in the same quintile among the children's distribution. For the 1990 cohort, while immobility at the bottom decreases, persistence at the top remains exactly the same.

Other cells of the transition matrices inform us about directional mobility: cells on the right-hand side of the diagonal represent upward mobility, while left-hand side cells represent downward mobility. It is important to note that directional mobility is not uniform; individuals in different areas of the distribution experience different mobility. For example, children of parents in Q2 in both cohorts are more likely to fall to Q1 than to move up to Q3. However, for children of parents in Q4, moving up to Q5 is slightly more likely than moving down to Q3. Additionally, for both cohorts, the share of individuals jumping from Q1 to Q5 is higher than the share of individuals falling from Q5 to Q1, but very few individuals experience these upward or downward leaps.

To better examine the trends of these transitional probabilities, I follow each *cell* throughout rolling birth cohorts. Figure 2 shows eight statistics divided into four panels: persistence at top/bottom, upward/downward from middle, upward/downward mobility (short movement), and extreme upward/downward mobility (long movement) as introduced in the previous section. All these transitional probabilities come from transitional matrices of each rolling cohort.

I will discuss two types of comparisons from the results in Figure 2: a comparison of the same statistics across cohorts and a comparison of the gap between two statistics (upward vs. downward) across cohorts. First, although there are different trends in these transitional probabilities, these changes are not substantial in magnitude. Most of the trends are either not statistically significant or are only weakly so. Table A2 reports the point estimates and

[10]While comparing the estimates to other studies is not the goal of this paper, an R-R of 0.6 is generally considered high in the literature.

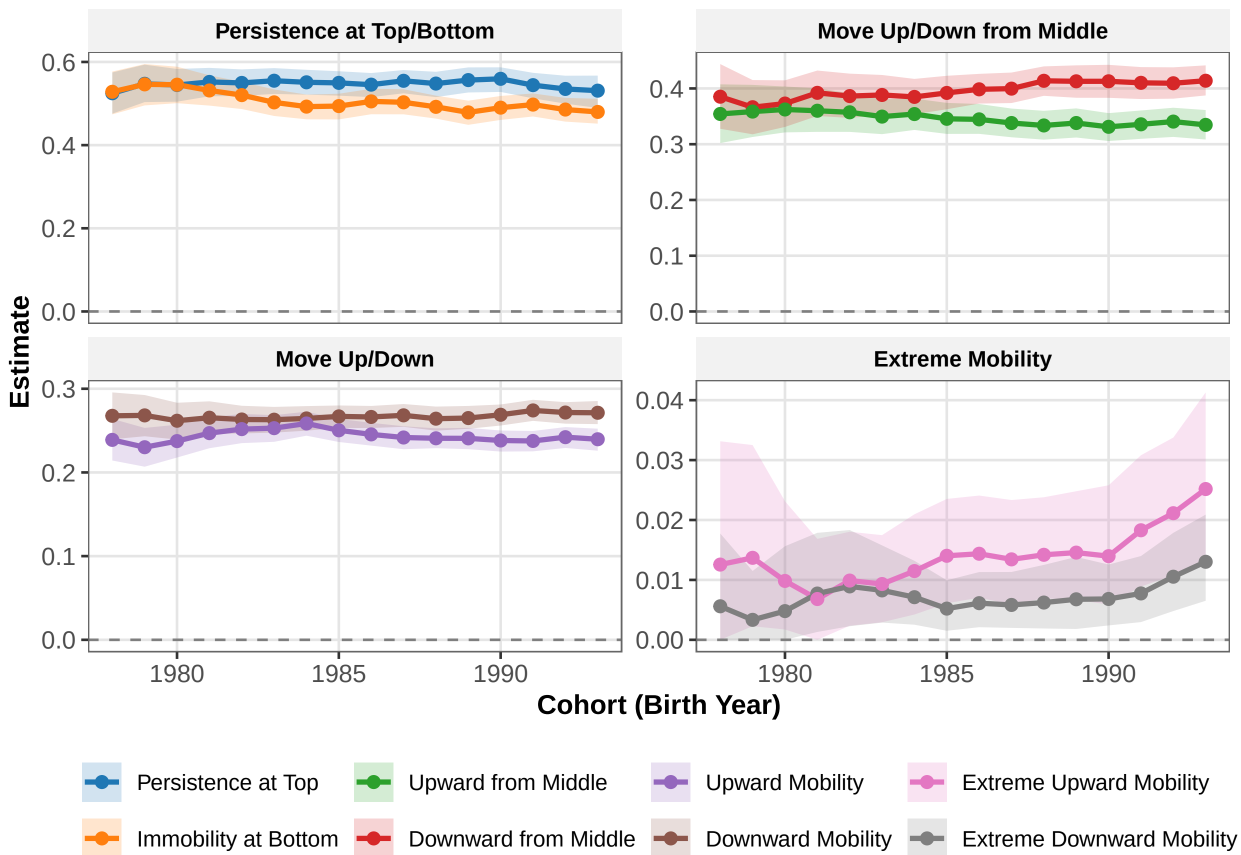

Figure 2: Trends in Directional Intergenerational Earnings Mobility

Notes: Earnings used in these panels are corrected via a Heckman-selection model. Shaded areas are 95% bootstrap confidence intervals with 1000 iterations.

SEs used to plot Figure 2. Immobility at the bottom decreases from 0.546 for the 1979 cohort to 0.480 for the 1993 cohort ($SE \approx 0.031; p \approx 0.031$). Downward mobility from middle increases from 0.366 for 1979 cohort to 0.414 for 1993 cohort ($SE \approx 0.029; p \approx 0.094$). Upward mobility decreases from 0.259 for 1984 cohort to 0.238 for 1991 cohort ($SE \approx 0.011; p \approx 0.048$). Although the point estimates for extreme upward mobility increase, there is no clear statistically significant trend. These results indicate that recent cohorts have more of a chance to break from the bottom quintile and reach the top quintile compared to previous cohorts. However, across the distribution, children are slightly less likely to have earnings one quintile higher than their parents. For those from a middle-earning family, they are even marginally more likely to fall to a lower quintile.

Second, there are significant gaps between the upward and downward statistics, and these gaps are more visible for children born after 1985. For the 1980 cohort, there is no difference between the share of individuals who are persistent at the top versus those who are immobile at the bottom (both are at 0.545). However, there is a gap of 0.069 ($SE \approx 0.021; p \approx 0.001$) for the 1990 cohort, driven by the reduction in immobility at the bottom. Similarly, the gap between upward mobility and downward mobility for children of middle-earning parents rises from 0.009 for the 1980 cohort to 0.082 ($SE \approx 0.020; p \approx 0.000$) for the 1990 cohort. The gap between total upward mobility and total downward mobility also rises from 0.025 ($SE \approx 0.015; p \approx 0.093$) for the 1980 cohort to 0.031 ($SE \approx 0.010; p \approx 0.002$) for the 1990 cohort. The extreme upward-downward mobility gap also increases after 1985, although this trend is not statistically significant. In general, I see a pattern whereby, while the opportunities to rise up show signs of improving through less immobility at the bottom and slightly higher extreme upward mobility, children in recent cohorts face higher risks of falling down than chances to move up.

These results show that mobility is very different for individuals at different segments of the earnings distribution. They also support my previous narrative about children having more compressed earnings with few positional exchanges. Combining the three discussed facts, (1) unchanged persistence at the top, (2) more downward movement from the middle, and (3) less immobility at the bottom, we can see that more equal earnings among children are associated with minor signs of positional exchanges in the middle of the distribution. The absolute gap in earnings between those from the bottom quintile and others tends to shrink, compressing the distribution to the right (Figure A3). This compression allows for more positional exchanges; however, without any large cross-cohort earnings fluctuations, these exchanges are modest. This narrative fits with the efforts of the Vietnamese government to reduce poverty with multiple social protection policies (Hoang-Duc et al., 2024).

In addition to earnings, I also examine the intergenerational persistence of education. Figure A6 shows the IGEdu using years of school and highest obtained degree by cohort. Similar to the IGE, the IGEdu trend is also linear and downward sloping, which implies that children's education for recent cohorts is less dependent on parental education. The transition matrices in Figure A7 show that this independence comes mostly from changes in the low and medium educational levels. From the 1980 cohort to the 1990 cohort, children of low-educated parents are more likely to obtain a medium education, while children of medium-educated parents are much less likely to have a lower education level. For these two groups of parents, the chance that their children will have a high education level remains the same between the two cohorts. One puzzle from these transition matrices is that, among parents with a high education level, fewer of them have children who obtain a high education level.

4.2 Subgroup analyses

This section delves into the heterogeneity in the levels and trends of the intergenerational mobility statistics discussed above. I will compare sons with daughters, urban with rural areas, and the north with the south. Here, for brevity, I will only discuss some of the most significant results. Other results can be found in the appendix.

Sons vs. daughters. There are multiple reasons that can lead to differences in IM between sons and daughters. Black and Devereux (2011) argue that assortative mating and labour supply responses play a significant role. Women from high-income families are likely to marry high-earning men, and if their labour supply responses involve choosing fewer hours of work, this can lead to a lower measured IGE for women compared to men. Mayer and Lopoo (2004) note that marital status and its changes over time influence mobility. The decline in marriage rates can affect the elasticity for both sons and daughters, especially if parental income is a better predictor of a child's own earnings than of their spouse's earnings. Gender-specific transmission channels can also be at play, as Brandén et al. (2024) find that rising persistence among women (mothers-daughters) is a driving force behind the overall decline in mobility.

I examine gender heterogeneity in mobility by repeating the main analyses for sons and daughters separately. Figure 3 reports the IGE, R-R, IGC, and AUM for males and females. Generally, the decreasing national trend remains for all measures. In addition, daughters tend to have lower economic dependence on their parents than sons do. Recent cohorts also

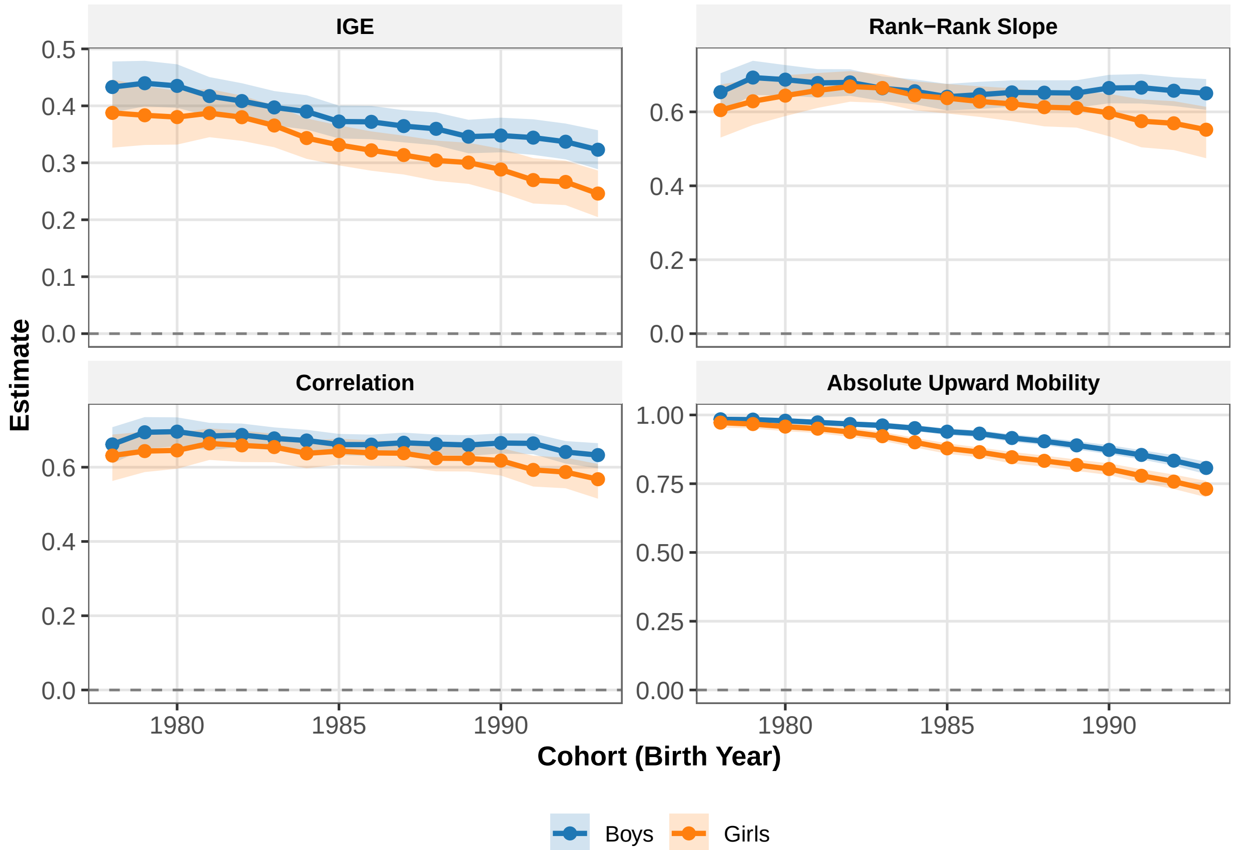

Figure 3: Trends in intergenerational earnings mobility between genders

Notes: Earnings used in these panels are corrected via a Heckman-selection model. Shaded areas are 95% bootstrap confidence intervals with 1000 iterations.

show sons having significantly higher persistence at the top than daughters (Figure A10). Part of this can be explained by social norms: in Vietnam, sons are still considered to be the main economic decision-makers and the heirs to continue the family lineage (Knodel et al., 2005); thus, it is not surprising that sons experience higher intergenerational persistence. It should also be noted that, due to the survey structure, it is difficult to distinguish between one's own children and in-laws. In Vietnam, it is a common practice for a married woman to live with her husband and his parents. Therefore, not all females in the sample are biological children. This can partially explain the lower level of intergenerational persistence for females compared to males.

Urban vs. rural. Differences in IM between urban and rural areas are influenced by several factors. Alesina et al. (2021) report that upward IM is generally higher in urban areas compared to rural areas. This is partly because agricultural employment is negatively correlated with upward mobility, while employment in services and manufacturing, which is more prevalent in urban settings, is positively correlated. Van der Weide et al. (2024) argue that greater access to resources and infrastructure, which is more prevalent in urban areas, is associated with higher societal mobility.

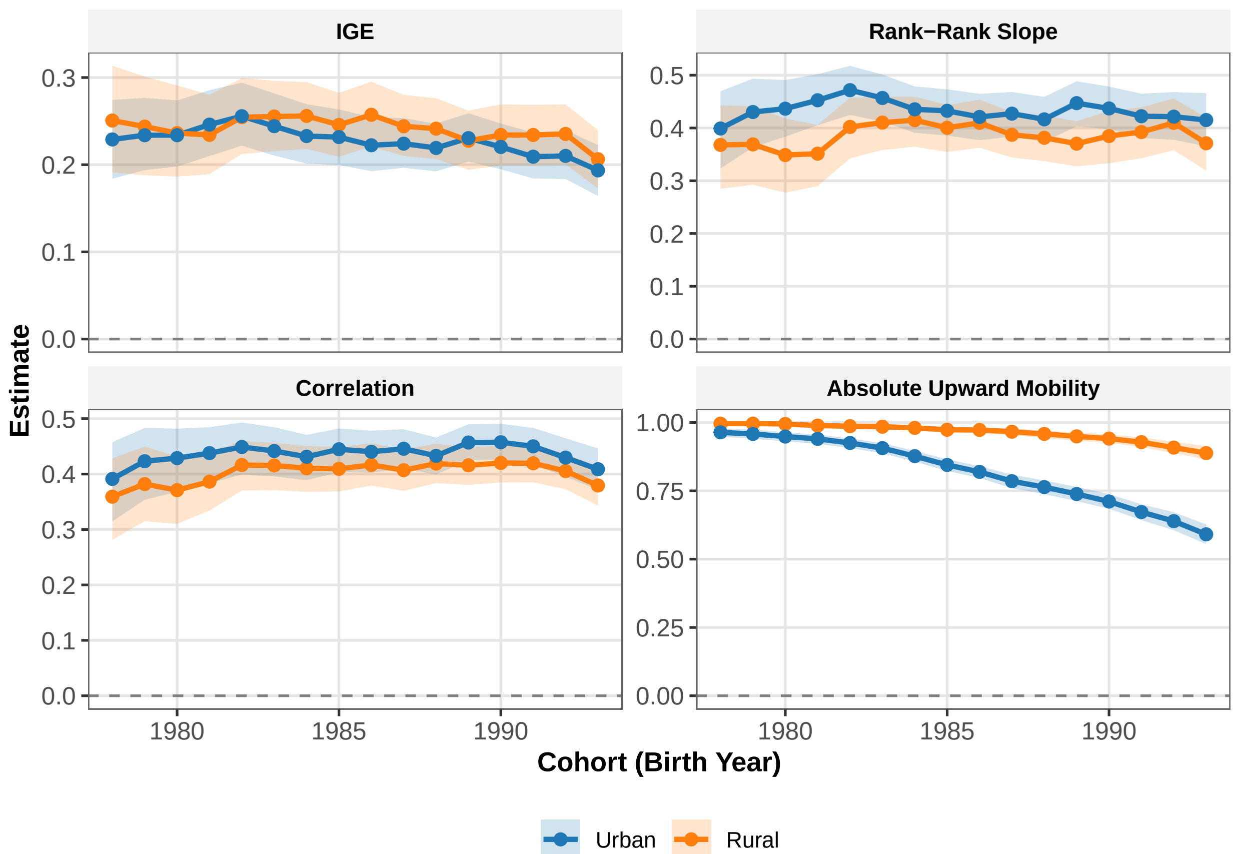

Figure 4: Trends in intergenerational earnings mobility between areas

Notes: Earnings used in these panels are corrected via a Heckman-selection model. Shaded areas are 95% bootstrap confidence intervals with 1000 iterations.

Figure 4 shows the main intergenerational mobility statistics for urban and rural areas. There are three interesting findings. First, for each subgroup, whether urban or rural, the mobility trends tend to be muted and do not follow the national trends. This signals that the national mobility trend is driven by gaps between rural and urban areas. Once that gap is controlled for, the IGE seems to be stable. Second, though not statistically significant, the IGC and R-R among urban families are consistently higher than those among rural families. Third, the AUM in urban areas falls much faster than the rural or national AUM. These last two findings suggest that the rate of earnings growth might differ between urban and rural areas. Over cohorts, parental earnings might outgrow children's earnings, causing the AUM trend to curve down more quickly for urban areas. That is less of the case in rural areas, which are characterised by a higher average child earnings growth rate. As rural children keep up with the earnings of their parents and of urban children, the left tail of the distribution is compressed, as I discussed earlier, reducing national inequality among children. To empirically test this hypothesis, I examine the ratio of earnings between urban and rural children. Figure A13 shows that this ratio is decreasing over cohorts, confirming this narrative.

North vs. South. Lastly, I examine the differences in mobility trends between northern and southern provinces.[11] In the literature, geographical differences in intergenerational mobility can arise from geography, historical legacies, or spatial sorting (Alesina et al., 2021). Throughout history, factors such as past investments, proximity to centers of economic activity, or terrain ruggedness can affect individuals' decisions to work and study, which in turn affects mobility. Other socioeconomic factors such as residential segregation, income inequality, the quality of primary schools, and social capital can also influence mobility (Chetty et al., 2014a).

Figure 5 reports the main intergenerational mobility statistics for northern and southern provinces. The national trends hold for these subgroups: the IGE decreases significantly, while the R-R and IGC are almost flat with marginal declines. The most important result from this comparison is that families in southern provinces experience significantly stronger intergenerational persistence than northern families. This seems counterintuitive, as the

[11] By northern provinces, I include all provinces in four out of the eight Economic Regions of Vietnam as categorised before 2023. This includes the Northeastern, Northwestern, Red River Delta, and North Central regions (which include Thanh Hoa, Nghe An, Ha Tinh, Quang Binh, Quang Tri, and Thua Thien Hue). All southern provinces are from the other four Economic Regions, including the South Central Coast, Central Highlands, Southeastern Region, and Mekong Delta.

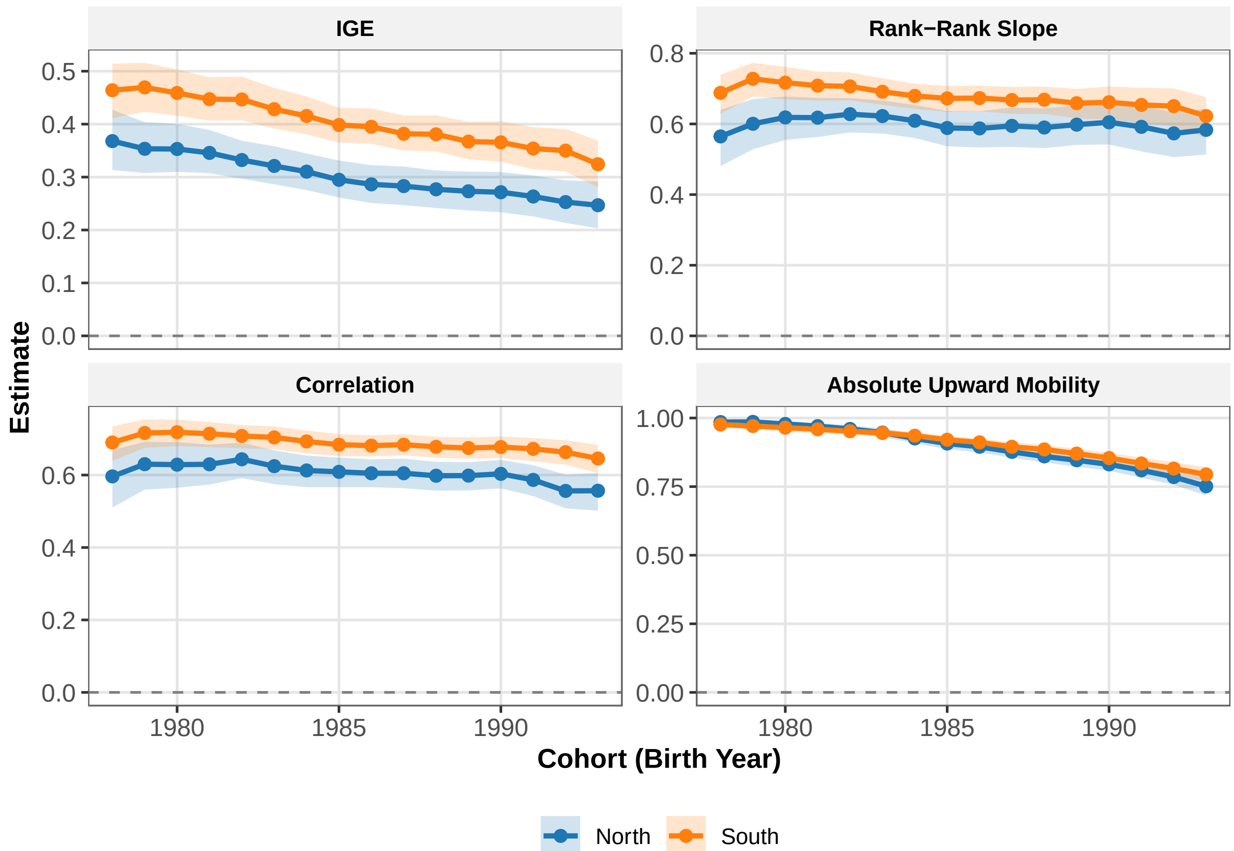

Figure 5: Trends in intergenerational earnings mobility between regions

Notes: Earnings used in these panels are corrected via a Heckman-selection model. Shaded areas are 95% bootstrap confidence intervals with 1000 iterations.

south of Vietnam is associated with opportunity and individualism (Nguyen, 2023; Ho et al., 2022), characteristics normally associated with mobility. However, the higher persistence makes sense when we consider inequality: southern provinces are more unequal than northern provinces (Figure A14). Also, families in the southern region experience much higher persistence at the top than those in the north (Figure A12).

5 Discussions of Trends

5.1 Link to education

Education has been proven to be among the strongest drivers of intergenerational mobility (Black and Devereux, 2011). Theoretical models, pioneered by Becker and Tomes (1979), posit that parental investment in children's education is a core driver of intergenerational linkages and the transmission of earnings and human capital. This relationship can also be strengthened by credit constraints: families facing credit constraints may exhibit higher intergenerational elasticities of income, as additional income has a greater impact on investment in their offspring's education (Black and Devereux, 2011). A higher return to education is found to increase intergenerational income persistence (Davis and Mazumder, 2024), and this effect is even complementary with parental human capital (Becker et al., 2018). Previous papers also examine intergenerational persistence in education and have found that parental education highly influences a child's educational attainment (Hu and Qian, 2023). This relationship is also found to vary greatly across geographical regions (Van der Weide et al., 2024; Alesina et al., 2021).

For developing countries that have experienced educational expansion, like Vietnam, it is important to study the relationship between education and intergenerational mobility, especially when the effects have been found to be mixed. On the one hand, increasing access to education or increasing the years of compulsory schooling directly improves the educational attainment of children, thus improving educational mobility (Van der Weide et al., 2024). A higher educational foundation in one generation can also promote mobility in later generations, especially for less developed regions (Alesina et al., 2021). On the other hand, education is a powerful tool for parents to pass down their resources, increasing intergenerational persistence. Black and Devereux (2011) report a few cases where an expansion in higher education can disproportionately benefit children from higher-income or wealthier families. Fan et al. (2021) document that the expansion of higher education in China, coupled with rising educational costs, has been linked to increasing intergenerational income persistence. Nybom and Stuhler (2024) study an educational expansion in Sweden and find non-monotonic effects: the reform increased mobility for the generation directly after the reform but reduced mobility for the generation afterward.

Vietnam has also experienced educational expansion, characterised by increased enrollment across all levels, with nearly universal enrollment in primary and lower secondary levels and a dramatic rise in upper secondary enrollment (Dang and Glewwe, 2018). In my sample, years of schooling also increase gradually among children across cohorts. This is associated with the increasing trends in educational mobility (Figure A6 and Figure A7). Variations in education also explain part of the earnings IGE and R-R. This can be seen by controlling for education in the mobility equations. Figure A15 shows that after controlling for the education of both children and parents, the IGE and R-R levels drop significantly for all cohorts, and the trends are muted. Education can explain more than 50% of the IGE and up to 30% of the R-R.

To formally examine the education-mobility relationship, this section adopts the model presented in Blanden et al. (2007), who examined the mediating roles of cognitive skills, non-cognitive skills, educational attainment, and labour market attachment in decreasing mobility in the UK. In this paper, I focus on using education, in different measures, to explain the trends in mobility as measured by the IGE in Vietnam.

Let y_i^c and y_i^p be the child and parent earnings, the intergenerational persistence is estimated via this simple empirical model introduced in the earlier section:

$$y_i^c = \alpha_0 + \beta y_i^p + \alpha_1 age_i^c + \alpha_2 (age_i^c)^2 + \alpha_3 age_i^p + \alpha_4 (age_i^p)^2 + \epsilon_i.$$

Let $\tilde{y}_i^c, \tilde{y}_i^p$ be the residual after partialling out the quadratic age terms, we can rewrite Equation 6 here:

$$\tilde{y}_i^c = \beta \tilde{y}_i^p + \varepsilon_i.$$

Structurally, let the relationship between education and earnings be determined by this system:

$$e_i = \omega_1 + \gamma y_i^p + u_{i1}, \tag{11}$$

$$y_i^c = \omega_2 + \lambda e_i + \nu_1 age_i^c + \nu_2 (age_i^c)^2 + u_{i2}. \tag{12}$$

This means that child's education, e_i, is determined by parental earnings. γ can be considered the return on investment or how much parental earnings explain child education. Child earnings is then determined by his/her own education and age to mitigate life-cycle bias. λ is the return to education, or educational payoff when children enter the labour market. In this empirical version, I abstract away from modelling unobserved child ability or other intergenerational transmission as in Becker and Tomes (1979) to focus on education.

By combining Equation 11 and Equation 12, we have:

$$y_i^c = \lambda\omega_1 + \omega_2 + \lambda\gamma y_i^p + \nu_1 age_i^c + \nu_2 (age_i^c)^2 + \lambda u_{i1} + u_{i2}.$$

We can residualise this by partialling out the age terms, similar to what we did above, and transform this into:

$$\tilde{y}_i^c = \underbrace{\lambda\gamma}_{\beta} \tilde{y}_i^p + \underbrace{(\widetilde{\lambda u_{i1} + u_{i2}})}_{\varepsilon_i}.$$

Standard OLS derivation of Equation 6 using this expression of $\tilde{y}_i^c$ returns:

$$\beta = \lambda\gamma + \frac{cov(\widetilde{\lambda u_{i1} + u_{i2}}; y_i^p)}{var(y_i^p)}.$$

This means the IGE can be decomposed into the sum of two components: the product of the return to education and the influence of parental earnings on child education, and the unexplained parent-child transmission channel. Therefore, I can examine the education-earnings relationship to explain the trend in mobility.

Table 4: Education related measures

Cohort	RoE degree	RoE higher	RoI degree	RoI higher	Expl. degree	Expl. higher	IGE	Unexpl. degree	Unexpl. higher
1978	0.197 (0.006)	0.607 (0.023)	1.269 (0.081)	0.294 (0.023)	0.250 (0.018)	0.178 (0.016)	0.425 (0.021)	0.175 (0.028)	0.247 (0.026)
1979	0.197 (0.006)	0.600 (0.021)	1.302 (0.072)	0.309 (0.021)	0.257 (0.016)	0.185 (0.014)	0.423 (0.019)	0.166 (0.025)	0.238 (0.023)
1980	0.197 (0.006)	0.589 (0.020)	1.276 (0.063)	0.311 (0.017)	0.252 (0.014)	0.183 (0.012)	0.412 (0.017)	0.160 (0.023)	0.229 (0.021)
1981	0.195 (0.006)	0.563 (0.019)	1.262 (0.054)	0.318 (0.015)	0.247 (0.013)	0.179 (0.010)	0.402 (0.016)	0.155 (0.020)	0.223 (0.019)
1982	0.191 (0.005)	0.543 (0.018)	1.272 (0.048)	0.335 (0.012)	0.243 (0.011)	0.182 (0.009)	0.394 (0.015)	0.151 (0.019)	0.213 (0.017)
1983	0.187 (0.005)	0.511 (0.017)	1.230 (0.044)	0.324 (0.012)	0.230 (0.011)	0.166 (0.008)	0.379 (0.014)	0.150 (0.017)	0.214 (0.016)
1984	0.183 (0.005)	0.485 (0.017)	1.207 (0.041)	0.328 (0.011)	0.221 (0.010)	0.159 (0.008)	0.365 (0.013)	0.144 (0.016)	0.206 (0.015)
1985	0.178 (0.005)	0.465 (0.016)	1.156 (0.039)	0.318 (0.011)	0.205 (0.009)	0.148 (0.007)	0.350 (0.012)	0.144 (0.015)	0.202 (0.014)
1986	0.175 (0.005)	0.453 (0.016)	1.157 (0.037)	0.312 (0.011)	0.202 (0.009)	0.141 (0.007)	0.344 (0.012)	0.142 (0.015)	0.203 (0.014)
1987	0.176 (0.005)	0.444 (0.015)	1.109 (0.035)	0.293 (0.011)	0.195 (0.008)	0.130 (0.007)	0.337 (0.012)	0.142 (0.015)	0.207 (0.014)
1988	0.172 (0.005)	0.434 (0.015)	1.111 (0.035)	0.295 (0.011)	0.191 (0.008)	0.128 (0.007)	0.331 (0.013)	0.140 (0.015)	0.203 (0.014)
1989	0.171 (0.005)	0.427 (0.015)	1.082 (0.035)	0.283 (0.011)	0.185 (0.008)	0.121 (0.006)	0.323 (0.013)	0.138 (0.015)	0.202 (0.014)
1990	0.170 (0.006)	0.417 (0.015)	1.076 (0.036)	0.276 (0.011)	0.183 (0.009)	0.115 (0.006)	0.320 (0.014)	0.137 (0.016)	0.205 (0.015)
1991	0.169 (0.006)	0.408 (0.015)	1.045 (0.037)	0.267 (0.012)	0.176 (0.008)	0.109 (0.006)	0.312 (0.014)	0.135 (0.017)	0.203 (0.016)
1992	0.166 (0.006)	0.399 (0.016)	1.039 (0.037)	0.268 (0.012)	0.172 (0.008)	0.107 (0.006)	0.305 (0.015)	0.133 (0.017)	0.198 (0.016)
1993	0.164 (0.006)	0.392 (0.016)	0.996 (0.038)	0.252 (0.012)	0.164 (0.008)	0.099 (0.006)	0.289 (0.016)	0.125 (0.018)	0.190 (0.017)

Note:
Each year of birth is a 5-year rolling birth cohort. Bootstrap standard errors are in parentheses.

Let's first examine the return to education (RoE), λ. The first two columns in Table 4 report the RoE estimated for a degree and for higher education. Although the magnitudes differ, it is clear that across cohorts, education tends to have lower marginal returns. For the 1980 cohort, obtaining one extra higher degree would, on average, be associated with 19.7% higher earnings. In that same cohort, those having higher education qualifications have almost 60% higher earnings than those without. However, these returns are only 17% and 41.7% for the 1990 cohort, respectively. This decreasing trend in the RoE can be observed in almost all subgroups (Figure A16 and Figure A18).

The influence of parental earnings on child education (return on investment - RoI), γ, also decreases over cohorts. Columns 3 and 4 in Table 4 report the RoI estimated for a degree and higher education, showing decreasing figures. This means that for later cohorts, parents need to earn more to increase their children's education. For the 1980 cohort, on average, in order for a child to have an additional degree, parental earnings must increase by a factor of 2.19.[12] For higher education, a child of a parent with 10% higher earnings is around 3% more likely to obtain a higher education qualification.[13] For the 1990 cohort, additional degree is related to more than 2.5 times higher parental earnings. 10% increase in parental earnings only associates with 2.6% higher probability of children obtaining higher education qualifications. This decreasing trend in the RoI can also be observed in almost all subgroups (Figure A17 and Figure A19).

From the previous decomposition practice, the $\lambda\gamma$ product is considered the part of IGE explained by education. Columns (4) and (5) in Table 4 report these statistics. The part of IGE explained by degree variable ranges between 0.164 for 1993 cohort to 0.257 for 1979 cohort. These are consistently higher than the part explained by higher education variable, which ranges from 0.099 for 1993 cohort to 0.185 for 1979 cohort. This is because higher education is a dummy variable, thus, captures less variation than the degree variable.

As RoE and RoI both decrease over cohorts, the $\lambda\gamma$ product also decreases. From the 1978 cohort to the 1993 cohort, $\lambda\gamma_{degree}$ decreases by 0.086 ($SE \approx 0.020; p \approx 0.000$) while

[12] We have $\Delta e_i = \gamma \times \Delta \ln(y_i^p)$. Let $\Delta e_i = 1$, with $\gamma = 1.276$, we need $\Delta \ln(y_i^p) = 2.19$.

[13] $\Delta e_i = \gamma \times \Delta \ln(y_i^p)$. With $\gamma = 0.311$, 10% increase in y_i^p would translate into roughly $0.311 \times \ln(1.1) \approx 0.029$ increase in e_i.

$\lambda\gamma_{HE}$ decreases by 0.079 ($SE \approx 0.017; p \approx 0.000$). Across these cohorts, IGE falls by 0.136 ($SE \approx 0.026; p \approx 0.000$). Given that, changes in $\lambda\gamma_{degree}$ explain more than 63% and changes in $\lambda\gamma_{HE}$ explains more than 58% of the change in IGE between 1978 and 1993 cohorts. The last two columns in Table 4 report the part of IGE left unexplained by education. After factoring out the degree variable, IGE only decreases by 0.050 ($SE \approx 0.033; p \approx 0.130$) and after factoring out the higher education variable, IGE falls only by 0.057 ($SE \approx 0.031; p \approx 0.066$). This means that changes in the education-earnings relationship play a key role in the observed increase in intergenerational earnings mobility in Vietnam.

The narrative behind this finding can be as follows. Educational expansion in Vietnam has come not only with broader access but also with a decreasing average return to education. This can result from a larger supply of medium- and high-educated individuals to the labour market, a mismatch between acquired skills and labour market demands (Le and Nguyen, 2016), or it can be due to "enrolment rising faster than quality" (Johns et al., 2007). Although Vietnamese students perform extremely well in international standardised tests such as PISA given the GDP per capita level, this performance is not uniform across all students from different backgrounds. While the influence of family resources on primary and secondary school enrollment has weakened, Dang and Glewwe (2018) report that there was a widening gap in higher-level educational outcomes between poor and better-off households. This aligns with this paper's findings so far: while the RoI is decreasing across cohorts, we can still see that children's education depends strongly on parental earnings. Figure 6 shows the mean of children's education per parental earnings rank. Children of the richest parents can obtain an additional three degrees and have a 60% higher chance of finishing any higher education qualifications than children of the poorest parents.

5.2 Link to inequality

Although inequality has been mentioned throughout the paper to reconcile the mobility findings, I devote this section to a concrete summary of the mobility-inequality relationship in Vietnam. It must be noted that the measure of inequality in this paper is different from the common approach. Most studies measure inequality at a specific point in time, considering

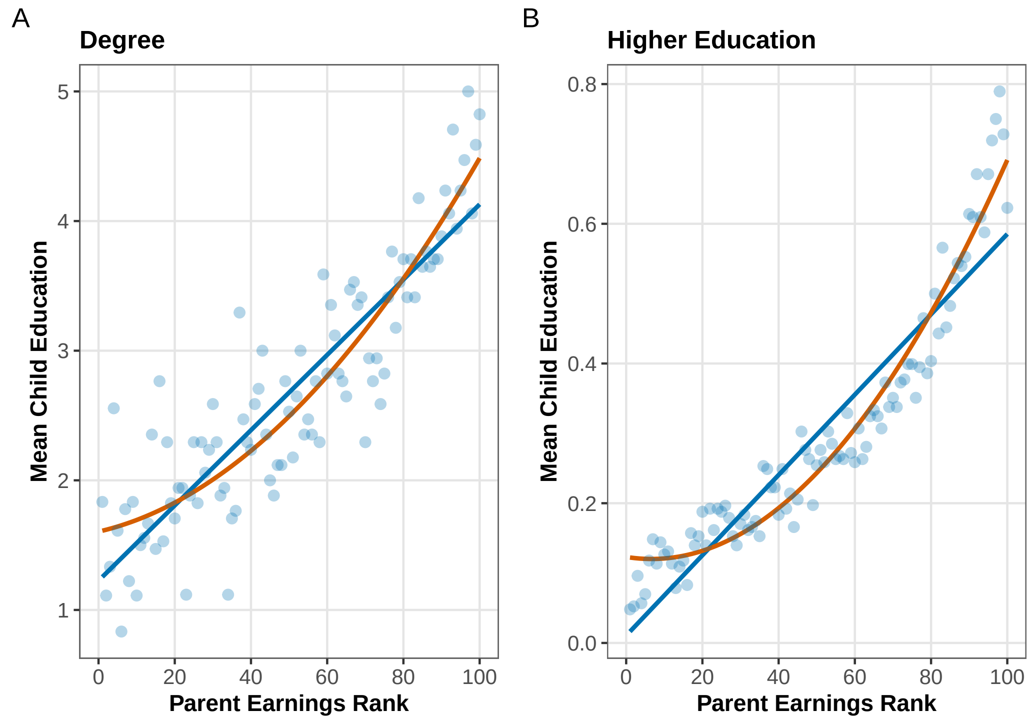

Figure 6: Child education and parental earnings rank

Notes: Earnings used in these panels are corrected via a Heckman-selection model. Higher education means any post-secondary formal degrees, including college diploma, bachelors, masters, and doctorate. Each subplot contains one linear fit and one quadratic fit.

the outcomes of different individuals at that time. For example, the earnings inequality of Vietnam in 2025 would be calculated using the earnings data of all surveyed Vietnamese in 2025, regardless of their characteristics: a 23-year-old can be compared to a 51-year-old. Changes in inequality are also interpreted across calendar years. This paper, on the other hand, measures inequality within each rolling child birth cohort. This means that I compare individuals who were born around the same year, whose earnings are measured in different VHLSS waves. Inequality changes in this paper are interpreted across cohorts. This practice will potentially yield different results from normal inequality reports. However, it aligns with the mobility trends analysis across cohorts.

The cross-sectional negative relationship between mobility and inequality is termed the "Great Gatsby Curve" (GGC) (Corak, 2013). This means that, on average, higher inequality in one generation is associated with lower economic mobility for the next generation.

The GGC has been observed across countries (Munoz and Van der Weide, 2025; Durlauf et al., 2022) and within countries (Chetty et al., 2014a; Deutscher and Mazumder, 2023). However, there are also cases where the GGC disappears or even reverses among developing countries (Genicot et al., 2024). Similarly, evidence on the relationship between mobility trends and inequality trends is mixed. Chetty et al. (2014b) establish that the R-R in the US is stable while inequality is rising. However, Durlauf et al. (2022) suggest that a decline in intergenerational mobility is associated with increasing inequality.

In Vietnam, trends in intergenerational mobility are largely driven by trends in inequality. I have reported three cases where this is observed. First, the decrease in the IGC, which is the IGE adjusted for inequality changes, is much smaller than the decrease in the IGE. This has been shown to be because, for recent cohorts, while parental earnings inequality remains stable, children's earnings inequality decreases (Figure 7). Second, the decreases in all relative non-directional mobility measures are muted after one considers the gap between rural and urban areas. This is due to the fact that children's earnings in rural areas grow faster than children's earnings in urban areas, closing the gap between the two areas (Figure A13). These two arguments are related by the fact that the earnings distribution is more compressed for later cohorts due to the left tail catching up (Figure A3). Finally, the North-South difference in mobility is associated with the difference in inequality: southern provinces are reported to have higher inequality and lower mobility than northern provinces. This aligns with the cross-sectional negative association of the GGC.

6 Conclusion

This paper has examined the trends in intergenerational earnings and education mobility in Vietnam for cohorts born between 1978 and 1993. Using five waves of the VHLSS between 2010 and 2018, a national cross-sectional household survey, I show that intergenerational mobility has been increasing for recent cohorts, and this finding holds across multiple relative non-directional measures, though the magnitudes differ. This increase is characterised by the "catching up" of children at the left tail of the earnings distribution. However, on average,

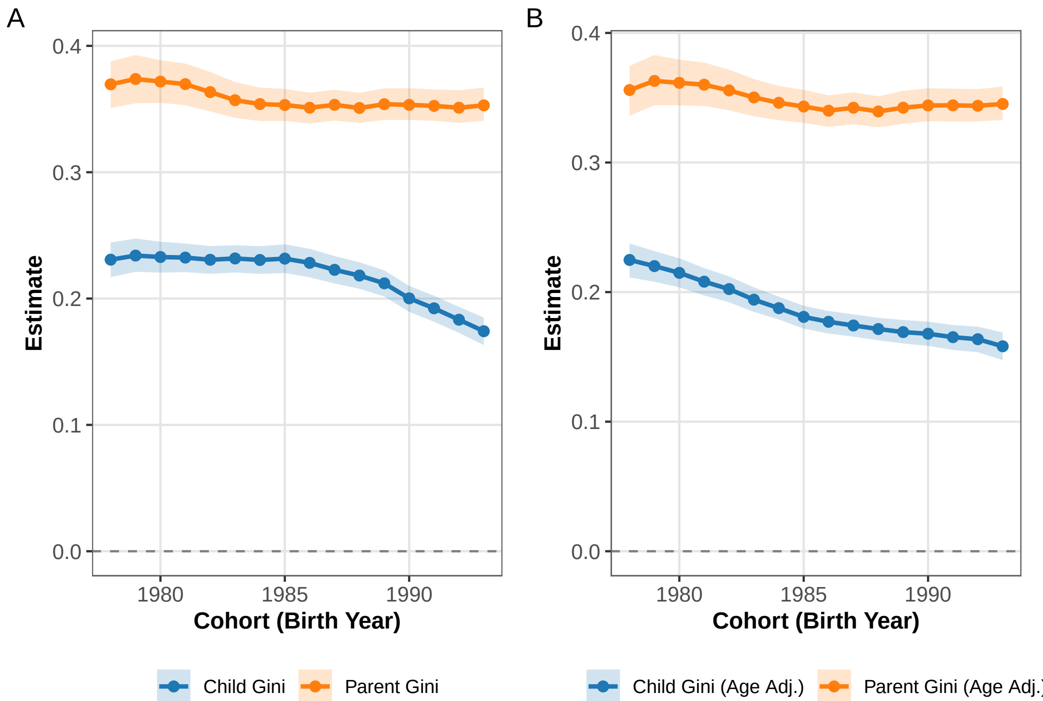

Figure 7: Trends in earnings inequality

Notes: Earnings used in these panels are corrected via a Heckman-selection model. Shaded areas are 95% bootstrap confidence intervals with 1000 iterations. 'Age Adj.' (Age adjusted) Gini coefficients are Gini calculated on the residuals of regressing earnings on age and age square.

children are becoming less likely to earn more than their parents.

Moreover, I implemented several subgroup analyses and found that sons' earnings tend to be more dependent on their parents than daughters', and that there can be differences in levels and trends between rural-urban areas and north-south regions. The evidence demonstrates that the narrowing rural-urban earnings gap in recent cohorts drives the national mobility trend. Additionally, southern provinces tend to exhibit higher intergenerational persistence than northern provinces.

Education and inequality play important roles in driving the reported trends. Changes in the relationship between earnings and education following educational expansion explain roughly 60% of the increase in earnings mobility. Changes in the ratio between the earnings inequality of parents and children are also found to alter the mobility trends measured by the IGE and IGC: when inequality is adjusted for, the increasing trends subside. The differences in mobility between rural and urban areas or the northern and southern regions are also explained partly by differences in inequality.

To mitigate empirical challenges, I adopted multiple appropriate methods and techniques, including a Heckman selection model, the rolling cohort setting, and age-adjusted earnings. However, due to data limitations, I could not fully address the challenges, and the estimates are likely biased. Due to a lack of detailed information about WFA individuals, or even missing information for those who moved out of their parents' house for more than 10 years, my Heckman correction procedure cannot fully address coresidency bias. Likewise, due to the cross-sectional structure, adjusting for age cannot comprehensively solve the life-cycle problem. Given these limitations, comparisons of mobility levels between Vietnam and other countries must be made with caution. This motivated my study of trends in the first place, instead of levels.

Going forward, more work needs to be done to reconcile the mobility estimates with the inequality trends, especially by using a wider set of data to study inequality among young people. intergenerational mobility is a multidimensional problem that interconnects with other socioeconomic topics; therefore, these mobility trends need to be compared to other trends, for example, in industrial development, structural transformation, migration, and urbanisation. One interesting topic for future research is to further examine the trends and gaps in mobility between rural and urban areas and their determinants.

References

Alesina, Alberto, Sebastian Hohmann, Stelios Michalopoulos, and Elias Papaioannou, "Intergenerational mobility in Africa," *Econometrica*, 2021, *89* (1), 1–35.

Becker, Gary S and Nigel Tomes, "An equilibrium theory of the distribution of income and intergenerational mobility," *Journal of Political Economy*, 1979, *87* (6), 1153–1189.

_ , Scott Duke Kominers, Kevin M Murphy, and Jörg L Spenkuch, "A theory of intergenerational mobility," *Journal of Political Economy*, 2018, *126* (S1), S7–S25.

Black, Sandra E and Paul J Devereux, "Recent developments in intergenerational mobility," in "Handbook of Labor Economics," Vol. 4, Elsevier, 2011, pp. 1487–1541.

Blanden, Jo, Alissa Goodman, Paul Gregg, and Stephen Machin, "Changes in intergenerational mobility in Britain," in "Generational income mobility in North America and Europe," Cambridge University Press, 2004, pp. 122–146.

_ , Paul Gregg, and Lindsey Macmillan, "Accounting for intergenerational income persistence: noncognitive skills, ability and education," *The Economic Journal*, 2007, *117* (519), C43–C60.

Brandén, Gunnar, Martin Nybom, and Kelly Vosters, "Like Mother, Like Child? The Rise of Women's Intergenerational Income Persistence in Sweden and the United States," *Journal of Labor Economics*, 2024.

Charles, Kerwin Kofi and Erik Hurst, "The correlation of wealth across generations," *Journal of Political Economy*, 2003, *111* (6), 1155–1182.

Chetty, Raj, David Grusky, Maximilian Hell, Nathaniel Hendren, Robert Manduca, and Jimmy Narang, "The fading American dream: Trends in absolute income mobility since 1940," *Science*, 2017, *356* (6336), 398–406.

_ , Nathaniel Hendren, Patrick Kline, and Emmanuel Saez, "Where is the land of opportunity? The geography of intergenerational mobility in the United States," *The Quarterly Journal of Economics*, 2014, *129* (4), 1553–1623.

_ , _ , _ , _ , and Nicholas Turner, "Is the United States still a land of opportunity? Recent trends in intergenerational mobility," *American Economic Review*, 2014, *104* (5), 141–147.

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

Dang, Hai-Anh H and Paul W Glewwe, "Well begun, but aiming higher: A review of Vietnam's education trends in the past 20 years and emerging challenges," *The Journal of Development Studies*, 2018, *54* (7), 1171–1195.

Dang, Thang, "Intergenerational earnings and income mobility in Vietnam," *Labour*, 2020, *34* (1), 113–136.

Davis, Jonathan MV and Bhashkar Mazumder, "The decline in intergenerational mobility after 1980," *Review of Economics and Statistics*, 2024, pp. 1–47.

der Weide, Roy Van, Christoph Lakner, Daniel Gerszon Mahler, Ambar Narayan, and Rakesh Gupta, "Intergenerational mobility around the world: A new database," *Journal of Development Economics*, 2024, *166*, 103167.

Deutscher, Nathan and Bhashkar Mazumder, "Measuring intergenerational income mobility: A synthesis of approaches," *Journal of Economic Literature*, 2023, *61* (3), 988–1036.

Doan, Quang Hung and Ngoc Anh Nguyen, "Intergenerational income mobility in Vietnam," *MPRA Paper No. 70603*, 2016, pp. 1–31.

Durlauf, Steven N, Andros Kourtellos, and Chih Ming Tan, "The Great Gatsby curve," *Annual Review of Economics*, 2022, *14* (1), 571–605.

Emran, M Shahe and Forhad Shilpi, "Economic approach to intergenerational mobility: Measures, methods, and challenges in developing countries," in "Social mobility in developing countries," Oxford University Press, 2021, pp. 197–220.

_ , William Greene, and Forhad Shilpi, "When measure matters: Coresidency, truncation bias, and intergenerational mobility in developing countries," *Journal of Human Resources*, 2018, *53* (3), 589–607.

Fan, Yi, Junjian Yi, and Junsen Zhang, "Rising intergenerational income persistence in China," *American Economic Journal: Economic Policy*, 2021, *13* (1), 202–230.

Fields, Gary S and Efe A Ok, "The measurement of income mobility: an introduction to the literature," *Handbook of Income Inequality Measurement*, 1999, pp. 557–598.

Genicot, Garance, Debraj Ray, and Carolina Concha-Arriagada, "Upward mobility in developing countries," Technical Report, National Bureau of Economic Research 2024.

Haider, Steven and Gary Solon, "Life-cycle variation in the association between current and lifetime earnings," *American Economic Review*, 2006, *96* (4), 1308–1320.

Heckman, James J, "Sample selection bias as a specification error," *Econometrica: Journal of the econometric society*, 1979, pp. 153–161.

Ho, Hoang-Anh, Peter Martinsson, and Ola Olsson, "The origins of cultural divergence: evidence from Vietnam," *Journal of Economic Growth*, 2022, *27* (1), 45–89.

Hoang-Duc, Chinh, Hang Nguyen-Thu, Tuan Nguyen-Anh, Hiep Tran-Duc, Linh Nguyen-Thi-Thuy, Phuong Do-Hoang, Nguyen To-The, Vuong Vu-Tien, and Huong Nguyen-Thi-Lan, "Governmental support and multidimensional poverty alleviation: efficiency assessment in rural areas of Vietnam," *The Journal of Economic Inequality*, 2024, *22* (4), 999–1038.

Hu, Yang and Yue Qian, "Gender, education expansion and intergenerational educational mobility around the world," *Nature Human Behaviour*, 2023, *7* (4), 583–595.

Iversen, Vegard, Anirudh Krishna, and Kunal Sen, *Social Mobility in Developing Countries: Concepts, Methods, and Determinants*, Oxford University Press, 2021.

Johns, Nicola, Minna Lyytikainen, Thi Thanh Thuy Nguyen, and Thi Lan Pham, "Education for All in Vietnam: High Enrolment, but Problems of Quality Remain," Technical Report, Young Lives Policy Brief 2007.

Knodel, John, Vu Manh Loi, Rukmalie Jayakody, and Vu Tuan Huy, "Gender roles in the family: Change and stability in Vietnam," *Asian Population Studies*, 2005, *1* (1), 69–92.

Le, Thuc Duc and Thi Thu Hang Nguyen, "Inequality in Educational Opportunities and Outcomes: Evidence From Young Lives Data in Vietnam," Technical Report, Young Lives Country Report 2016.

Lee, Chul-In and Gary Solon, "Trends in intergenerational income mobility," *The Review of Economics and Statistics*, 2009, *91* (4), 766–772.

Major, Lee Elliot and Stephen Machin, *Social mobility: And its enemies*, Penguin UK, 2018.

Mayer, Susan E and Leonard M Lopoo, "What do trends in the intergenerational economic mobility of sons and daughters in the United States mean," *Generational income mobility in North America and Europe*, 2004, pp. 90–121.

Munoz, Ercio and Roy Van der Weide, "Intergenerational Income Mobility around the World: A New Database," Technical Report, The World Bank Policy Research Working Papers 2025.

Nguyen, Bach, "Regional informal institutions, local governance and internal migration in Vietnam," *Regional Studies*, 2023, *57* (7), 1189–1206.

Nguyen, Cuong Viet and Lam Tran Nguyen, "Intra-generational and intergenerational social mobility: Evidence from Vietnam," *Journal of Asian and African Studies*, 2020, *55* (3), 370–397.

Nybom, Martin and Jan Stuhler, "Biases in standard measures of intergenerational income dependence," *Journal of Human Resources*, 2017, *52* (3), 800–825.

_ **and** _ , "Interpreting trends in intergenerational mobility," *Journal of Political Economy*, 2024, *132* (8), 2531–2570.

Solon, Gary, "Intergenerational mobility in the labor market," in "Handbook of Labor Economics," Vol. 3, Elsevier, 1999, pp. 1761–1800.

Appendix

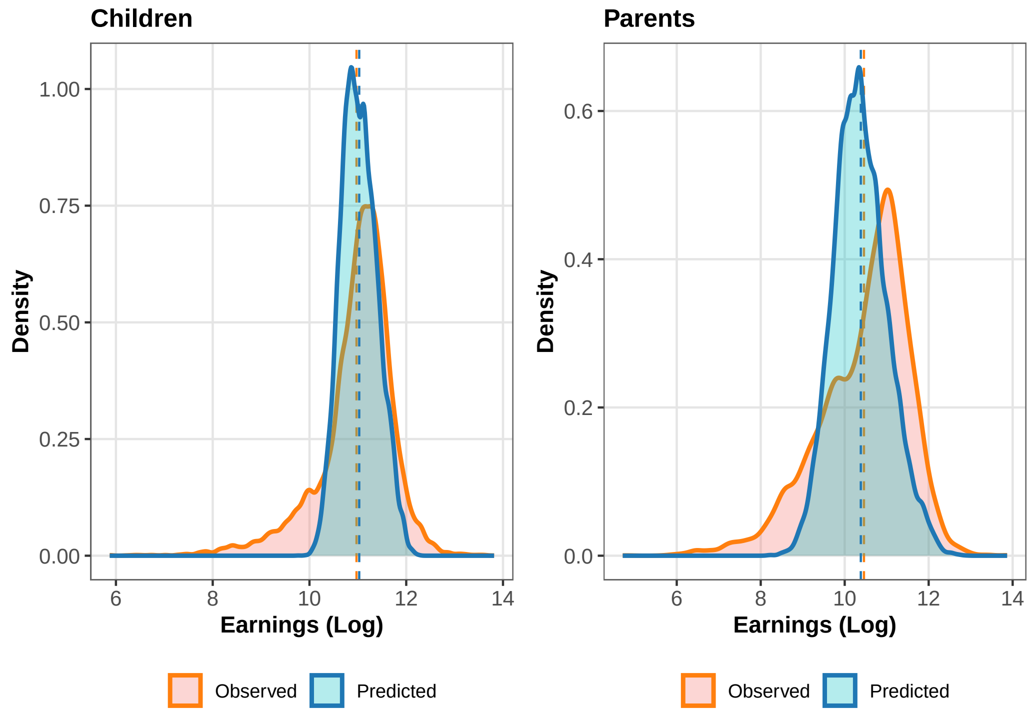

Figure A1: Distributions of observed and predicted earnings

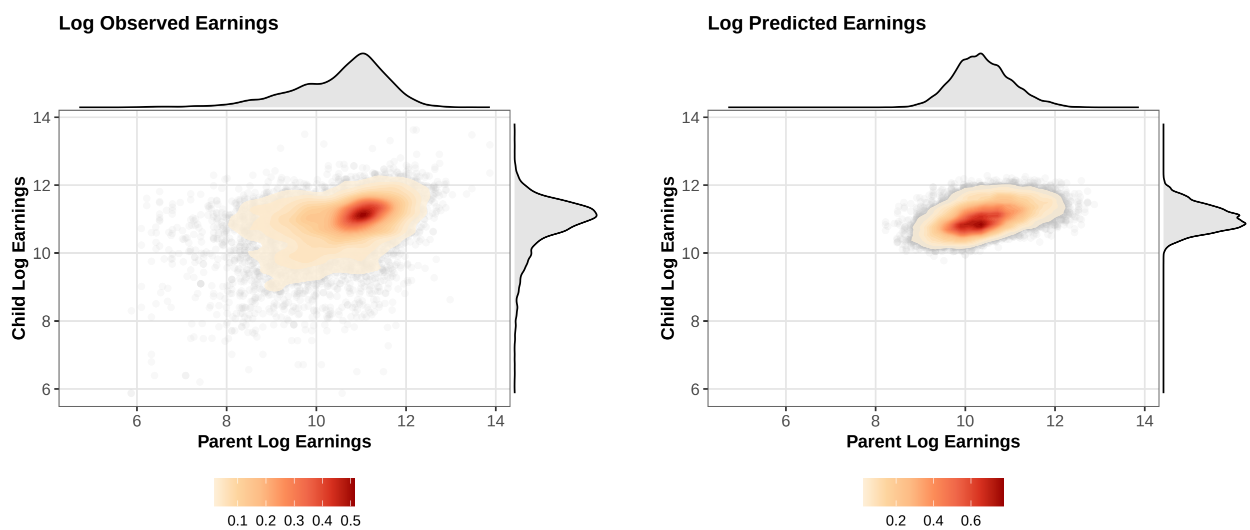

Figure A2: Joint distributions of child and parent earnings

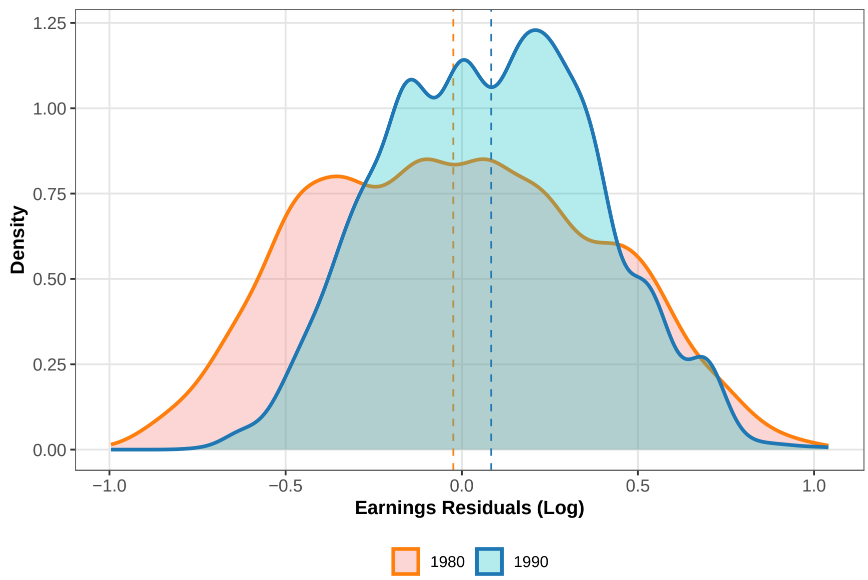

Figure A3: Distributions of age-adjusted earnings for two cohorts

Figure A4: Rank-rank slope of predicted earnings

Notes: Earnings used in these panels are predicted via the Heckman-selection model. The rank-rank slopes are obtained by regressing percentile rank of child predicted earnings on parent predicted earnings without age controls.

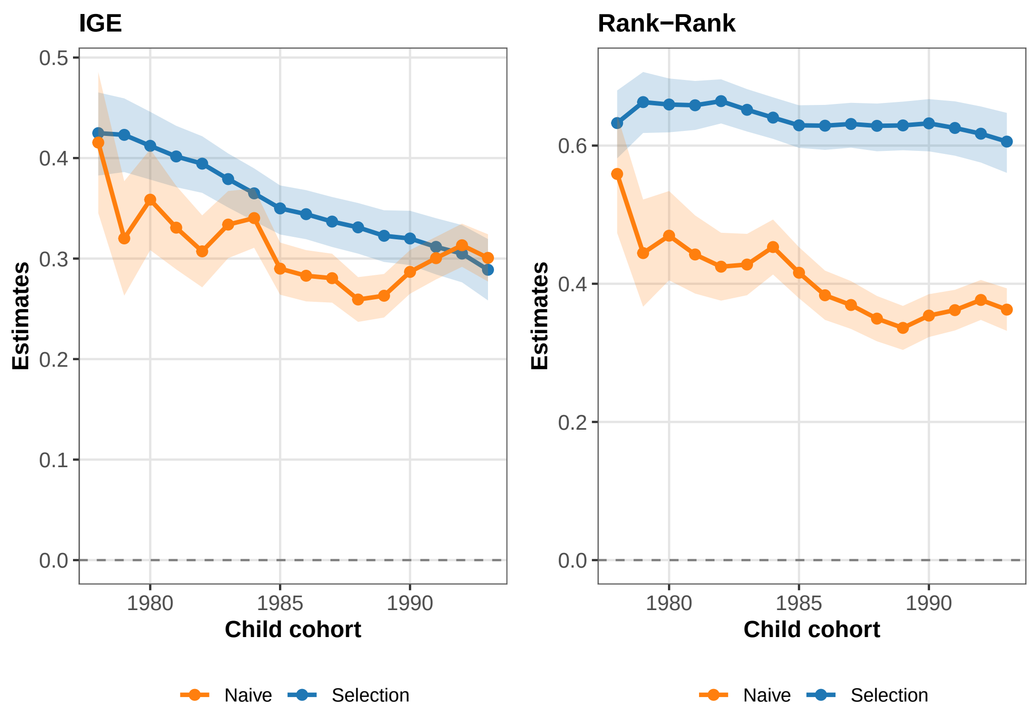

Figure A5: Trends in intergenerational earnings mobility: Naive vs Selection

Notes: Shaded areas for naive earnings are 95% confidence intervals. Shaded areas for selection-adjusted earnings are 95% bootstrap confidence intervals with 1000 iterations.

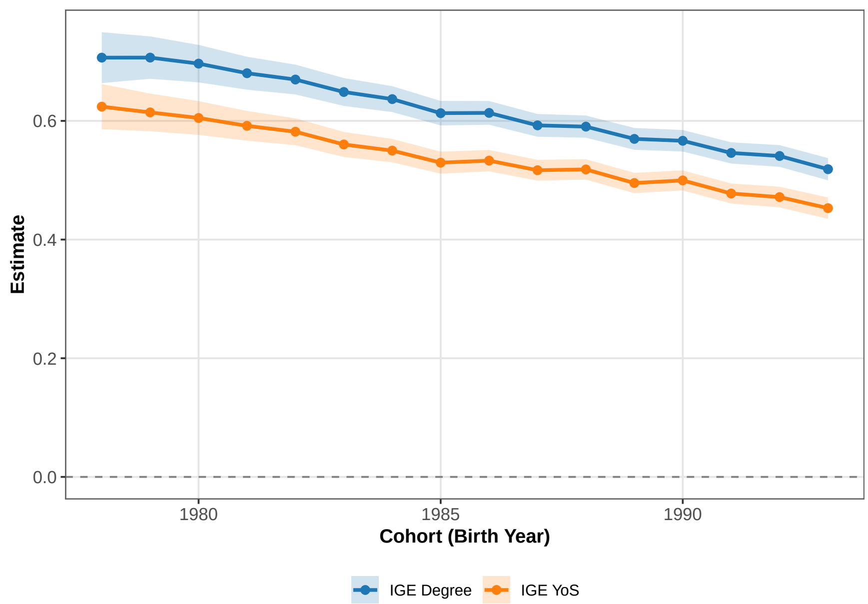

Figure A6: Trends in Intergenerational Education Mobility

Notes: Shaded areas are 95% confidence intervals. 'IGE Degree' is the estimate from regressing degree (8 levels: no degree, primary, secondary, high school, college diploma, bachelors, masters, doctorate) of children on that of their parents. 'IGE YoS' is the estimate from regressing year of schooling of children on that of their parents.

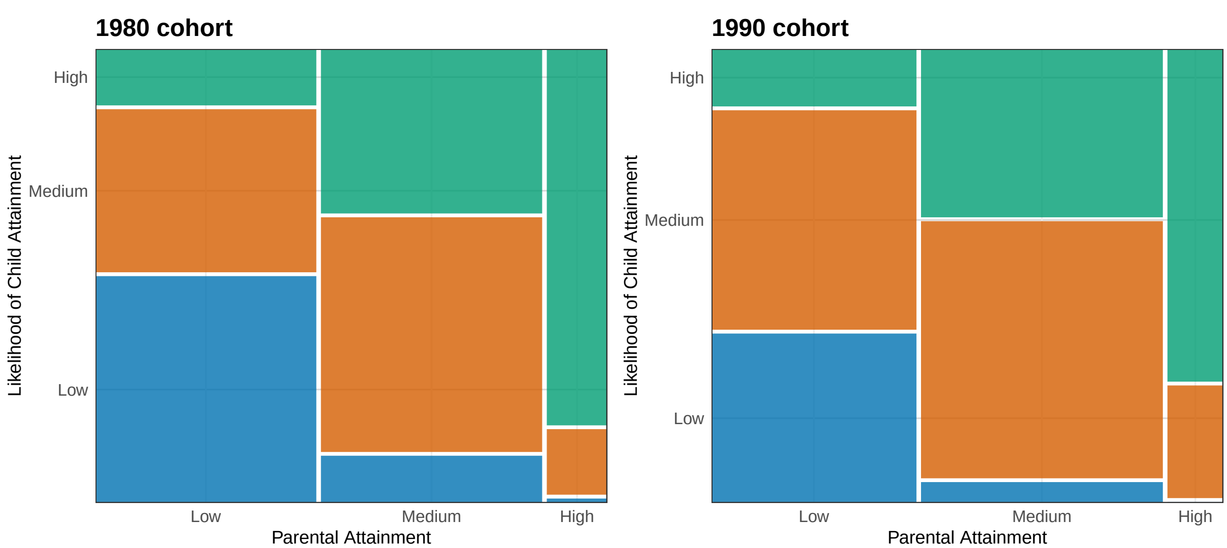

Figure A7: Intergenerational Education Transitional Matrix

Notes: 'Low' educational level means ones with no or primary certificate. 'Medium' level means ones with secondary and high school diploma. 'High' level means any higher post-secondary formal education.

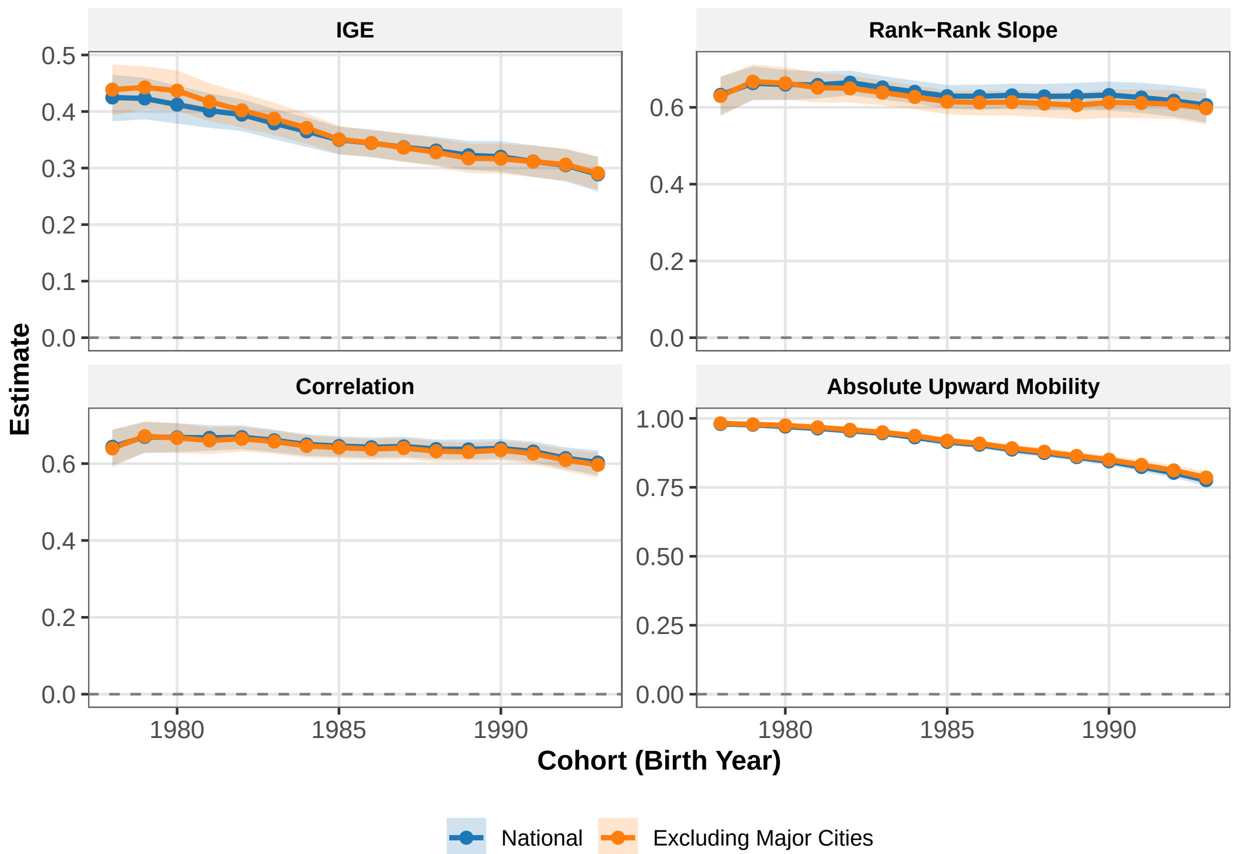

Figure A8: Trends in intergenerational earnings mobility: Dropping big cities

Notes: Earnings used in these panels are corrected via a Heckman-selection model. Shaded areas are 95% bootstrap confidence intervals with 1000 iterations. Hanoi and Ho Chi Minh cities are dropped from the sample.

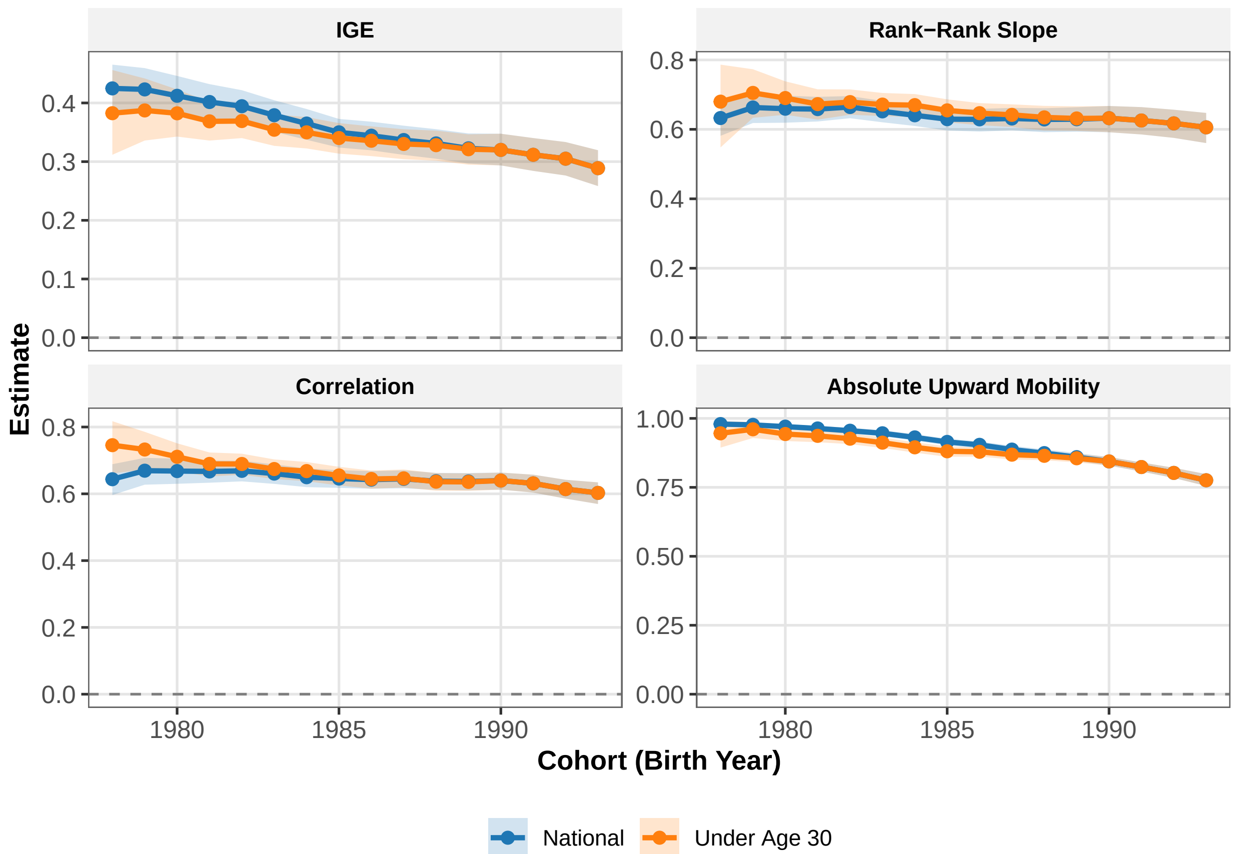

Figure A9: Trends in intergenerational earnings mobility: Young people

Notes: Earnings used in these panels are corrected via a Heckman-selection model. Shaded areas are 95% bootstrap confidence intervals with 1000 iterations. Only children under or equal 30 are kept in this sample.

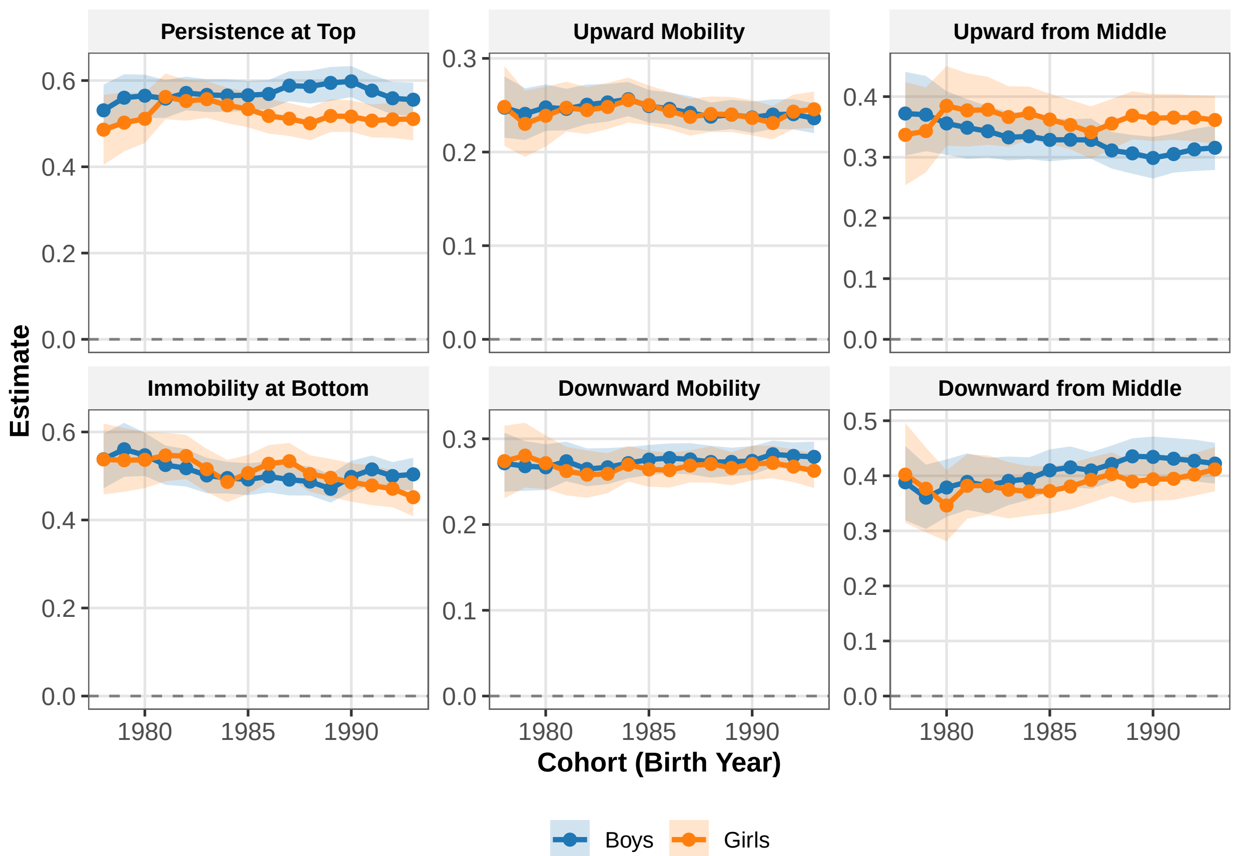

Figure A10: Trends in directional intergenerational earnings mobility between genders

Notes: Earnings used in these panels are corrected via a Heckman-selection model. Shaded areas are 95% bootstrap confidence intervals with 1000 iterations.

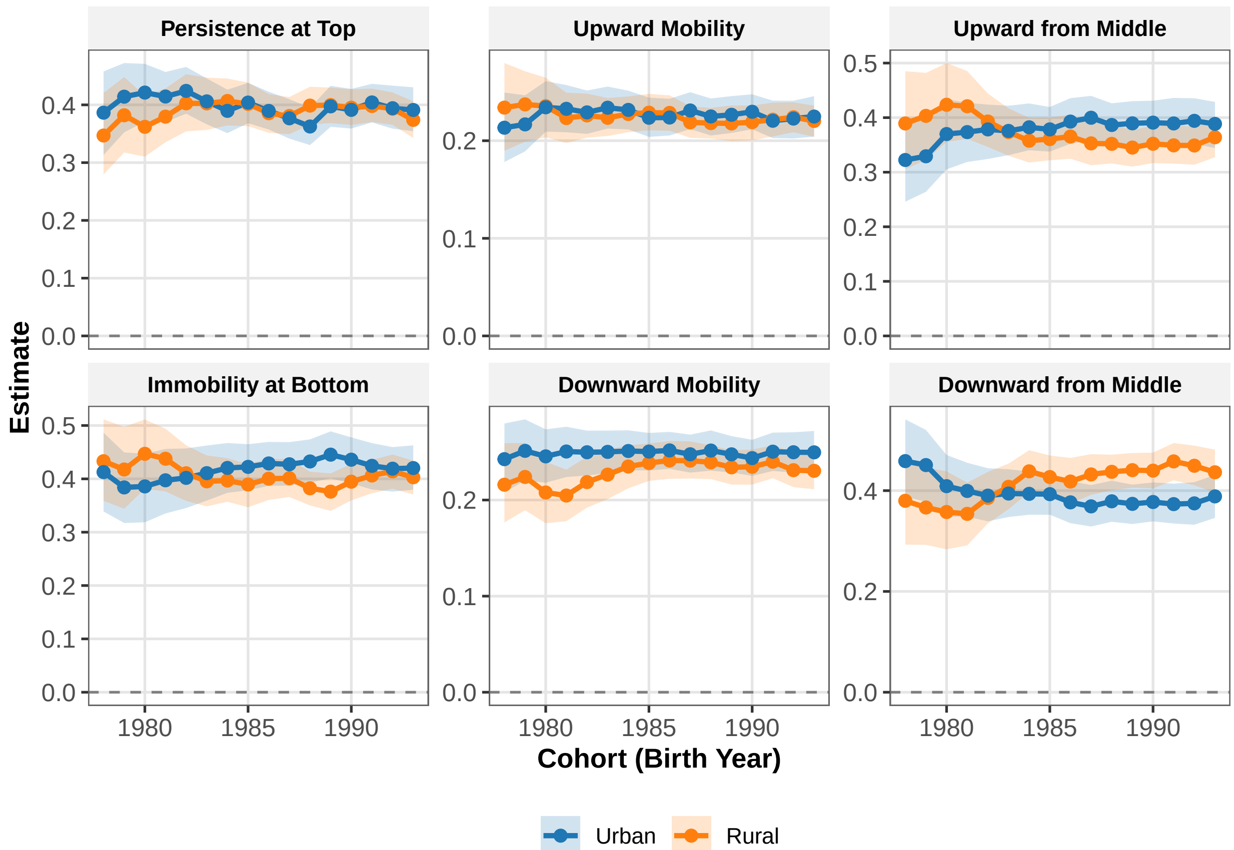

Figure A11: Trends in directional intergenerational earnings mobility between areas

Notes: Earnings used in these panels are corrected via a Heckman-selection model. Shaded areas are 95% bootstrap confidence intervals with 1000 iterations.

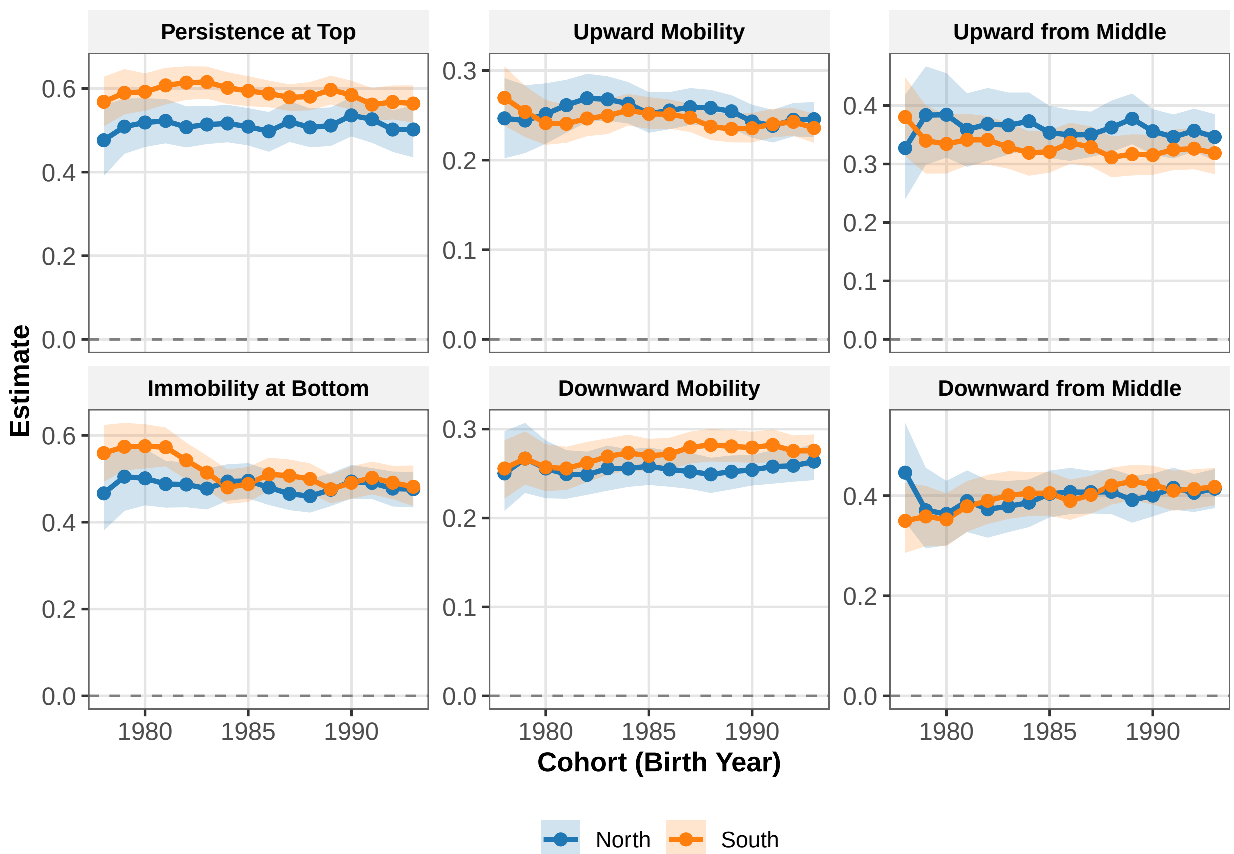

Figure A12: Trends in directional intergenerational earnings mobility between regions

Notes: Earnings used in these panels are corrected via a Heckman-selection model. Shaded areas are 95% bootstrap confidence intervals with 1000 iterations.

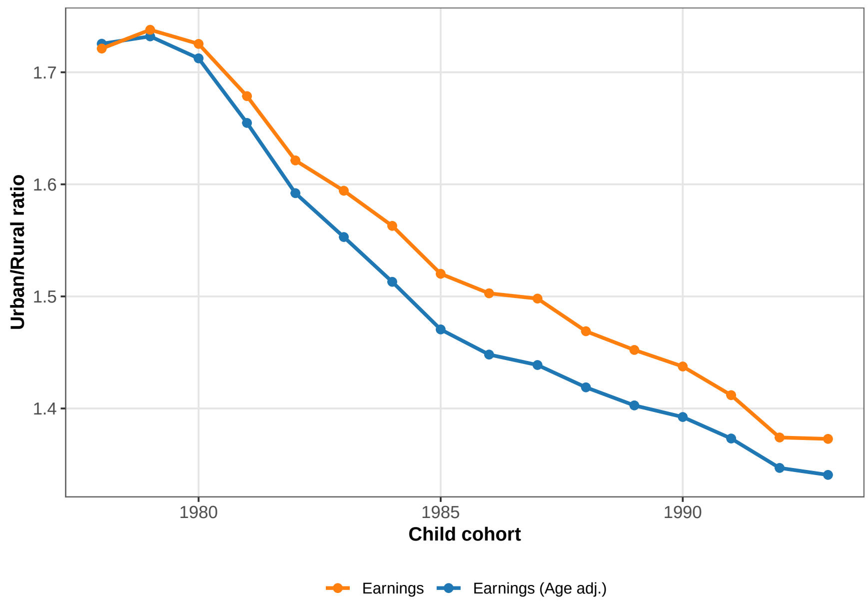

Figure A13: Urban/Rural ratio of children's earnings

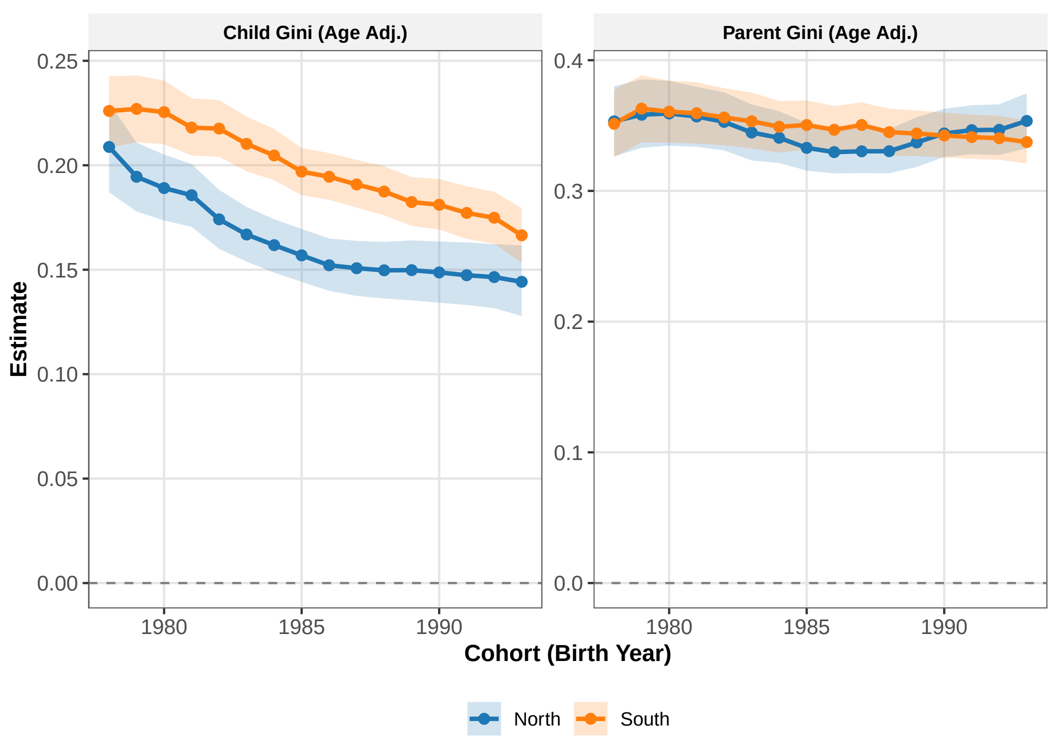

Figure A14: Trends in earnings inequality between regions

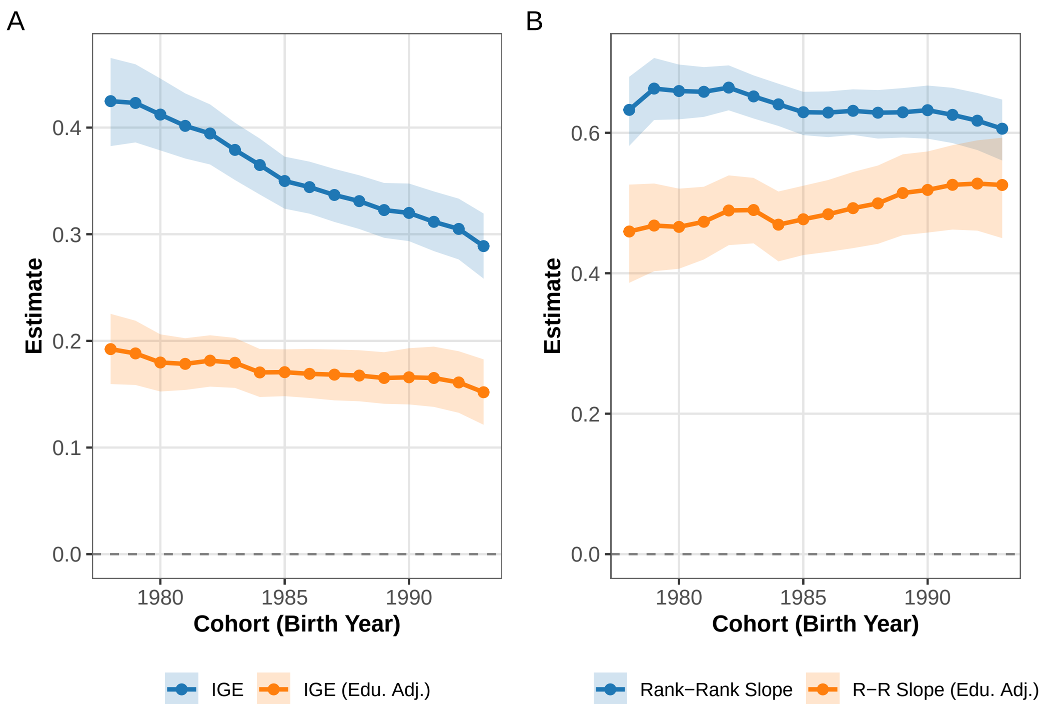

Figure A15: Trends in intergenerational earnings mobility: Control for education

Notes: Earnings used in these panels are corrected via a Heckman-selection model. Shaded areas are 95% bootstrap confidence intervals with 1000 iterations. Education variables included are degrees obtained by children and parents (8 levels).

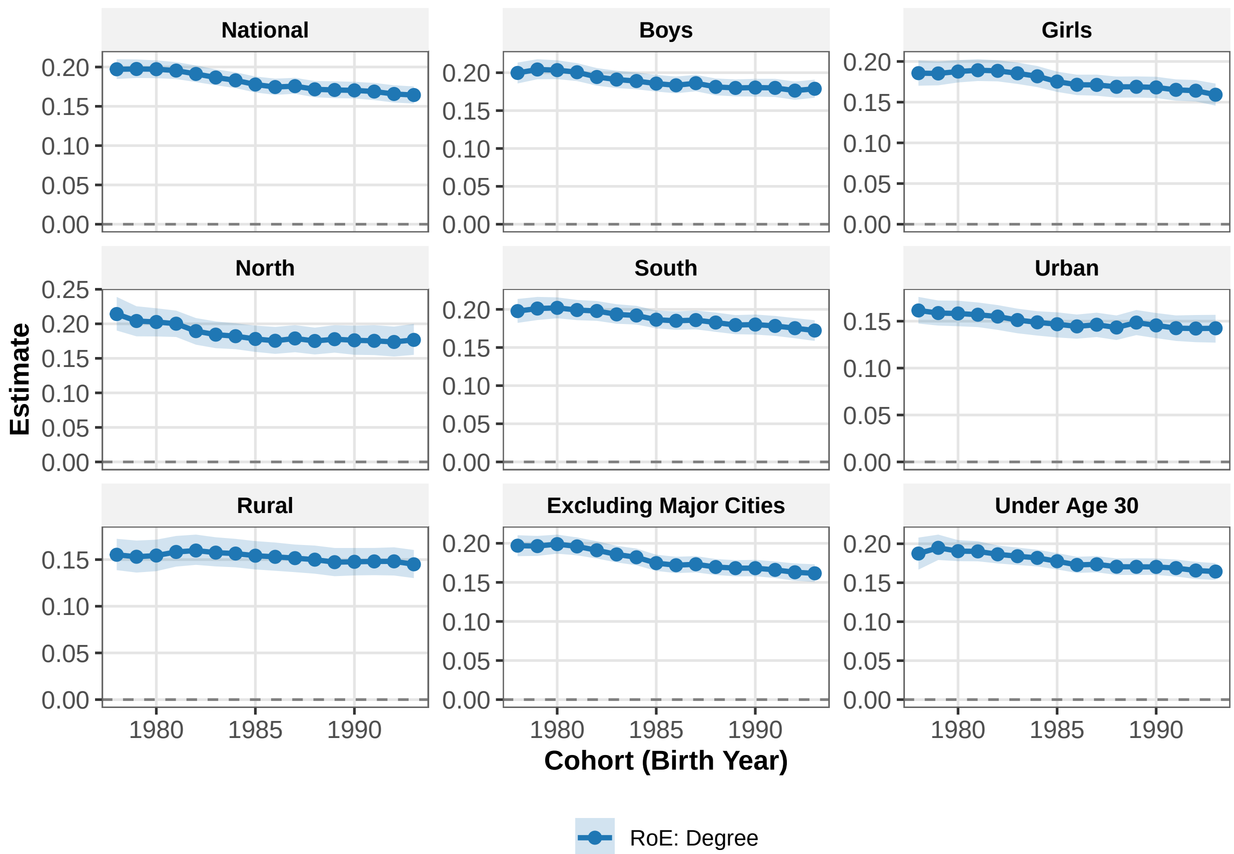

Figure A16: Trends in earnings returns to degree

Notes: Earnings used in these panels are corrected via a Heckman-selection model. Shaded areas are 95% bootstrap confidence intervals with 1000 iterations. Degree means the level of highest formal qualification.

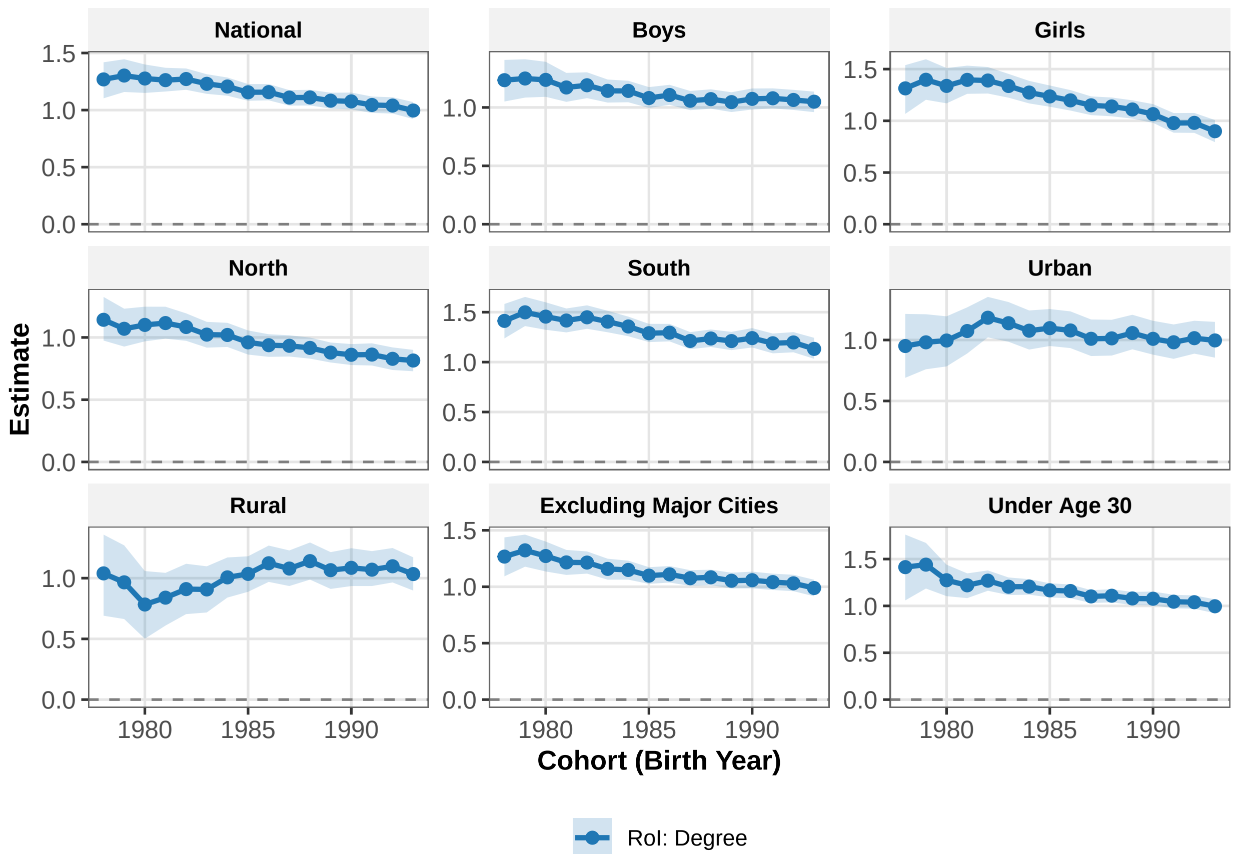

Figure A17: Trends in human capital returns to investment: Degree

Notes: Earnings used in these panels are corrected via a Heckman-selection model. Shaded areas are 95% bootstrap confidence intervals with 1000 iterations. Degree means the level of highest formal qualification.

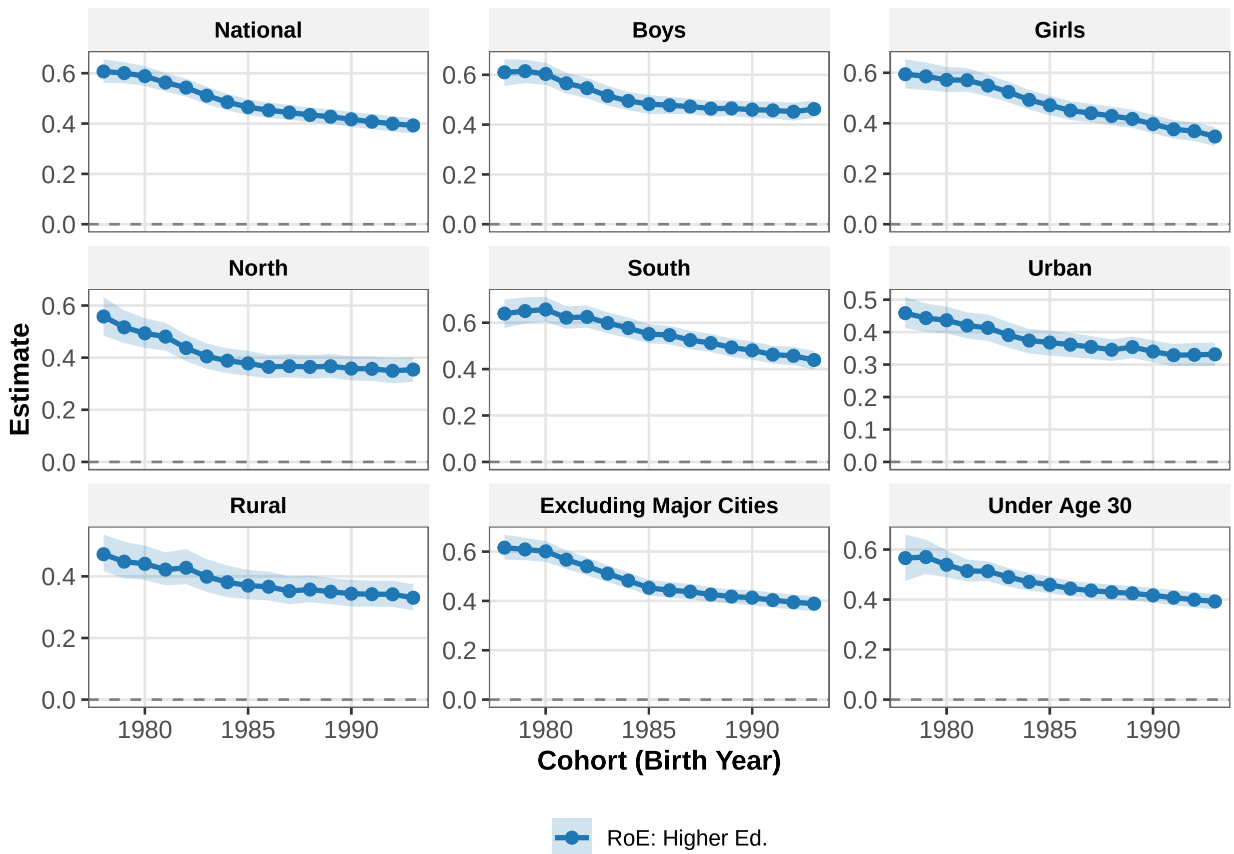

Figure A18: Trends in earnings returns to higher education

Notes: Earnings used in these panels are corrected via a Heckman-selection model. Shaded areas are 95% bootstrap confidence intervals with 1000 iterations. Higher education means any post-secondary formal degrees, including college diploma, bachelors, masters, and doctorate.

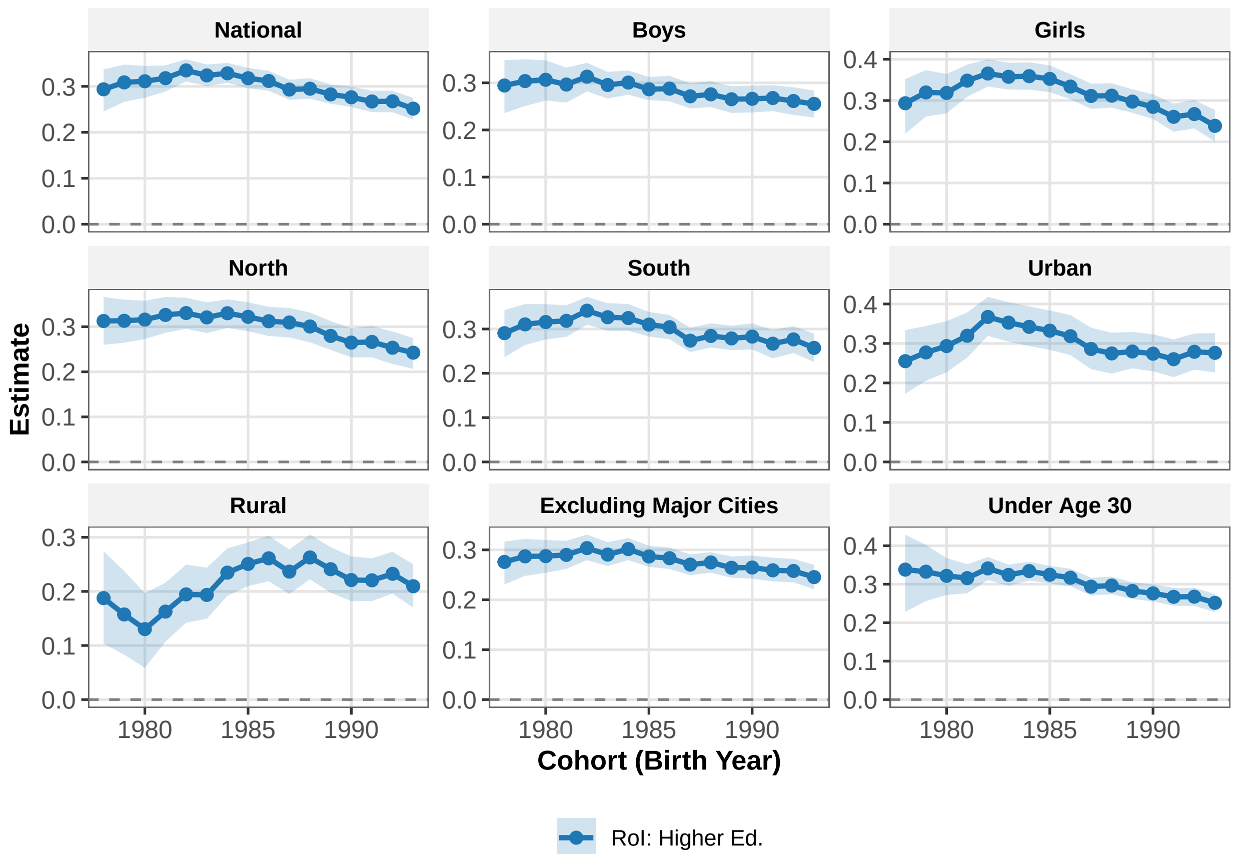

Figure A19: Trends in human capital returns to investment: Higher education

Notes: Earnings used in these panels are corrected via a Heckman-selection model. Shaded areas are 95% bootstrap confidence intervals with 1000 iterations. Higher education means any post-secondary formal degrees, including college diploma, bachelors, masters, and doctorate.

Table A1: Intergenerational mobility by cohorts

Cohort	IGE		R-R		Corr.		IGE ctrl.		R-R ctrl.		Corr. ctrl.		IGE ctrl. 2		R-R ctrl. 2		Corr. ctrl. 2	
1978	0.425	(0.021)	0.633	(0.026)	0.644	(0.024)	0.192	(0.017)	0.459	(0.034)	0.495	(0.033)	0.373	(0.039)	0.590	(0.047)	0.603	(0.047)
1979	0.423	(0.019)	0.663	(0.023)	0.669	(0.021)	0.188	(0.015)	0.468	(0.032)	0.525	(0.029)	0.346	(0.035)	0.567	(0.045)	0.603	(0.044)
1980	0.412	(0.017)	0.659	(0.021)	0.668	(0.019)	0.180	(0.014)	0.466	(0.029)	0.513	(0.028)	0.324	(0.031)	0.568	(0.041)	0.589	(0.042)
1981	0.402	(0.016)	0.658	(0.018)	0.667	(0.017)	0.178	(0.013)	0.473	(0.027)	0.515	(0.026)	0.316	(0.028)	0.576	(0.037)	0.597	(0.037)
1982	0.394	(0.015)	0.664	(0.016)	0.669	(0.016)	0.181	(0.012)	0.489	(0.025)	0.518	(0.025)	0.311	(0.025)	0.579	(0.035)	0.592	(0.035)
1983	0.379	(0.014)	0.652	(0.016)	0.661	(0.015)	0.179	(0.012)	0.490	(0.024)	0.514	(0.023)	0.301	(0.024)	0.574	(0.033)	0.588	(0.033)
1984	0.365	(0.013)	0.641	(0.015)	0.650	(0.014)	0.170	(0.011)	0.469	(0.025)	0.485	(0.024)	0.297	(0.022)	0.576	(0.032)	0.578	(0.031)
1985	0.350	(0.012)	0.629	(0.016)	0.646	(0.014)	0.171	(0.011)	0.477	(0.026)	0.495	(0.023)	0.297	(0.022)	0.588	(0.032)	0.592	(0.030)
1986	0.344	(0.012)	0.629	(0.016)	0.642	(0.013)	0.169	(0.012)	0.484	(0.027)	0.494	(0.023)	0.297	(0.021)	0.599	(0.032)	0.600	(0.029)
1987	0.337	(0.012)	0.631	(0.017)	0.645	(0.013)	0.168	(0.012)	0.493	(0.028)	0.503	(0.023)	0.295	(0.022)	0.609	(0.032)	0.612	(0.028)
1988	0.331	(0.013)	0.629	(0.017)	0.638	(0.013)	0.167	(0.012)	0.500	(0.029)	0.500	(0.022)	0.298	(0.022)	0.632	(0.033)	0.625	(0.027)
1989	0.323	(0.013)	0.629	(0.018)	0.637	(0.013)	0.165	(0.012)	0.514	(0.030)	0.506	(0.022)	0.296	(0.023)	0.634	(0.034)	0.627	(0.027)
1990	0.320	(0.014)	0.632	(0.020)	0.640	(0.013)	0.166	(0.013)	0.519	(0.031)	0.511	(0.022)	0.298	(0.024)	0.644	(0.034)	0.636	(0.026)
1991	0.312	(0.014)	0.626	(0.021)	0.631	(0.014)	0.165	(0.014)	0.526	(0.032)	0.498	(0.024)	0.298	(0.025)	0.649	(0.035)	0.633	(0.027)
1992	0.305	(0.015)	0.617	(0.021)	0.614	(0.014)	0.161	(0.015)	0.528	(0.033)	0.485	(0.025)	0.296	(0.027)	0.660	(0.036)	0.628	(0.029)
1993	0.289	(0.016)	0.606	(0.023)	0.603	(0.016)	0.152	(0.016)	0.526	(0.037)	0.474	(0.029)	0.286	(0.029)	0.657	(0.038)	0.619	(0.032)

Note:
Each year of birth is a 5-year rolling birth cohort. Bootstrap standard errors are in parentheses.

Table A2: Intergenerational mobility by cohorts (cont.)

Cohort	Pers. High	Pers. Low	Up from Mid	Down from Mid	Upward	Downward	Extr. Upward	Extr. Downward
1978	0.525 (0.026)	0.528 (0.027)	0.354 (0.027)	0.385 (0.029)	0.239 (0.013)	0.268 (0.014)	0.013 (0.008)	0.006 (0.005)
1979	0.548 (0.023)	0.546 (0.026)	0.359 (0.024)	0.366 (0.025)	0.230 (0.012)	0.268 (0.013)	0.014 (0.008)	0.003 (0.003)
1980	0.545 (0.020)	0.545 (0.022)	0.362 (0.021)	0.373 (0.022)	0.237 (0.010)	0.262 (0.011)	0.010 (0.006)	0.005 (0.004)
1981	0.552 (0.018)	0.531 (0.019)	0.360 (0.020)	0.392 (0.021)	0.247 (0.009)	0.265 (0.010)	0.007 (0.004)	0.008 (0.004)
1982	0.550 (0.016)	0.520 (0.017)	0.357 (0.018)	0.386 (0.020)	0.252 (0.009)	0.263 (0.009)	0.010 (0.004)	0.009 (0.004)
1983	0.555 (0.016)	0.503 (0.016)	0.350 (0.016)	0.388 (0.018)	0.253 (0.008)	0.263 (0.008)	0.009 (0.004)	0.008 (0.003)
1984	0.551 (0.015)	0.493 (0.015)	0.354 (0.015)	0.385 (0.016)	0.259 (0.008)	0.265 (0.007)	0.011 (0.004)	0.007 (0.003)
1985	0.550 (0.014)	0.494 (0.016)	0.346 (0.014)	0.392 (0.016)	0.250 (0.007)	0.267 (0.007)	0.014 (0.004)	0.005 (0.002)
1986	0.545 (0.014)	0.505 (0.016)	0.345 (0.014)	0.398 (0.014)	0.245 (0.007)	0.266 (0.007)	0.014 (0.004)	0.006 (0.002)
1987	0.555 (0.014)	0.503 (0.015)	0.338 (0.013)	0.400 (0.014)	0.242 (0.007)	0.268 (0.007)	0.013 (0.004)	0.006 (0.002)
1988	0.548 (0.015)	0.492 (0.014)	0.334 (0.014)	0.414 (0.014)	0.241 (0.006)	0.264 (0.007)	0.014 (0.004)	0.006 (0.003)
1989	0.556 (0.015)	0.479 (0.015)	0.338 (0.013)	0.413 (0.015)	0.241 (0.006)	0.265 (0.007)	0.015 (0.005)	0.007 (0.003)
1990	0.559 (0.015)	0.490 (0.015)	0.331 (0.013)	0.413 (0.015)	0.238 (0.007)	0.269 (0.007)	0.014 (0.005)	0.007 (0.003)
1991	0.544 (0.016)	0.497 (0.014)	0.336 (0.013)	0.410 (0.015)	0.238 (0.006)	0.274 (0.006)	0.018 (0.006)	0.008 (0.003)
1992	0.535 (0.017)	0.486 (0.015)	0.341 (0.013)	0.409 (0.014)	0.242 (0.006)	0.272 (0.007)	0.021 (0.006)	0.011 (0.003)
1993	0.531 (0.020)	0.480 (0.016)	0.335 (0.014)	0.414 (0.014)	0.240 (0.007)	0.271 (0.007)	0.025 (0.007)	0.013 (0.004)

Note:
Each year of birth is a 5-year rolling birth cohort. Bootstrap standard errors are in parentheses.

Table A3: First Stage Heckman Selection Models (Probit)

	Child stays home	Parent stays home	Child has job	Parent has job
Communal WFA Rate (Child)	−2.966***	−6.802***		
	(0.001)	(0.007)		
Children / HH Members			1.501***	−1.496***
			(0.007)	(0.005)
Education	−0.020***	0.062***	0.123***	−0.007***
	(0.001)	(0.004)	(0.001)	(0.001)
Vocational Education	0.313***	−4.279*	1.192***	−0.342***
	(0.003)	(2.486)	(0.005)	(0.004)
Female	0.335***	−0.437***	0.323***	−0.090***
	(0.002)	(0.007)	(0.003)	(0.002)
Age	−0.132***	0.599***	1.540***	0.083***
	(0.004)	(0.009)	(0.005)	(0.005)
Age^2 / 100	0.245***	−0.602***	−2.458***	−0.220***
	(0.006)	(0.008)	(0.008)	(0.004)
Urban	−2.109***	24.950***	11.843***	0.568***
	(0.089)	(0.540)	(0.098)	(0.182)
Region	−0.889***	2.095***	9.998***	1.116***
	(0.033)	(0.153)	(0.042)	(0.072)
Urban x Region	0.945***	−8.801***	−7.613***	−3.792***

Table A3 continued from previous page

	Child stays home	Parent stays home	Child has job	Parent has job
	(0.053)	(0.308)	(0.057)	(0.105)
Urban x Education	0.372***	0.084***	0.207***	0.206***
	(0.002)	(0.005)	(0.002)	(0.001)
Urban x Vocational	0.133***	3.836	−0.160***	−0.420***
	(0.005)	(2.486)	(0.007)	(0.005)
Urban x Female	−0.178***	0.983***	−0.183***	−0.638***
	(0.004)	(0.011)	(0.005)	(0.003)
Urban x Age	0.085***	−0.950***	−0.916***	−0.057***
	(0.006)	(0.020)	(0.007)	(0.007)
Urban x Age^2 / 100	−0.145***	0.874***	1.571***	0.047***
	(0.010)	(0.018)	(0.012)	(0.006)
Region x Education	−0.008***	−0.009***	−0.057***	0.027***
	(0.001)	(0.002)	(0.001)	(0.001)
Region x Vocational	−0.181***	3.806	−0.419***	0.264***
	(0.002)	(2.486)	(0.003)	(0.003)
Region x Female	−0.309***	−0.002	−0.483***	−0.398***
	(0.002)	(0.004)	(0.002)	(0.001)
Region x Age	0.090***	−0.095***	−0.642***	−0.085***
	(0.002)	(0.006)	(0.003)	(0.003)
Region x Age^2 / 100	−0.167***	0.104***	1.038***	0.110***

Table A3 continued from previous page

	Child stays home	Parent stays home	Child has job	Parent has job
	(0.004)	(0.005)	(0.005)	(0.002)
Urban x Region x Education	−0.183***	−0.048***	−0.059***	−0.109***
	(0.001)	(0.003)	(0.001)	(0.001)
Urban x Region x Vocational	0.082***	−3.520	0.087***	0.357***
	(0.003)	(2.486)	(0.004)	(0.003)
Urban x Region x Female	0.098***	−0.444***	0.189***	0.332***
	(0.002)	(0.007)	(0.003)	(0.002)
Urban x Region x Age	−0.029***	0.326***	0.556***	0.160***
	(0.004)	(0.011)	(0.004)	(0.004)
Urban x Region x Age^2 / 100	0.051***	−0.294***	−0.941***	−0.154***
	(0.006)	(0.011)	(0.007)	(0.004)
Year Fixed Effect: 2012	−0.968***	−0.449***	0.113***	0.101***
	(0.001)	(0.003)	(0.001)	(0.001)
Year Fixed Effect: 2014	−0.015***	−0.190***	0.120***	0.177***
	(0.001)	(0.003)	(0.001)	(0.001)
Year Fixed Effect: 2016	−0.041***	0.061***	0.245***	0.274***
	(0.001)	(0.003)	(0.001)	(0.001)
Year Fixed Effect: 2018	−0.086***	−0.085***	0.267***	0.250***
	(0.001)	(0.003)	(0.001)	(0.001)
Intercept	2.932***	−11.907***	−22.925***	3.103***

Table A3 continued from previous page

	Child stays home	Parent stays home	Child has job	Parent has job
	(0.055)	(0.237)	(0.070)	(0.126)
N.	20,649	13,958	17,795	22,599

* p < 0.1, ** p < 0.05, *** p < 0.01

Table A4: Second Stage Heckman Selection Models (OLS)

	Child Earnings (WFA)	Parent Earnings (WFA)	Child Earnings (Job)	Parent Earnings (Job)
Inverse Mills Ratio	−0.291***	0.325***	0.032	−0.013
	(0.023)	(0.104)	(0.145)	(0.118)
Education	0.203***	0.463***	0.206***	0.461***
	(0.021)	(0.031)	(0.021)	(0.031)
Vocational Education	0.209***	−0.078	0.249***	−0.046
	(0.064)	(0.107)	(0.074)	(0.107)
Female	−0.148***	−0.533***	−0.129**	−0.522***
	(0.055)	(0.066)	(0.059)	(0.067)
Age	0.351***	−0.087	0.374***	−0.104
	(0.083)	(0.106)	(0.109)	(0.111)

Table A4 continued from previous page

	Child WFA Selection	Parent WFA Selection	Child Job Selection	Parent Job Selection
Age^2 / 100	−0.544***	0.031	−0.580***	0.048
	(0.142)	(0.100)	(0.182)	(0.106)
Urban	1.746	−19.568***	1.683	−20.097***
	(1.629)	(5.160)	(1.691)	(5.161)
Region	1.830***	−2.717	2.022**	−2.775
	(0.708)	(1.748)	(0.840)	(1.779)
Urban x Region	−1.116	8.640***	−1.119	8.823***
	(0.956)	(3.094)	(1.008)	(3.106)
Urban x Education	−0.081***	−0.057	−0.049	−0.060
	(0.030)	(0.043)	(0.033)	(0.045)
Urban x Vocational	0.033	0.204	0.033	0.187
	(0.092)	(0.148)	(0.093)	(0.150)
Urban x Female	0.026	0.175	0.021	0.156
	(0.075)	(0.111)	(0.076)	(0.125)
Urban x Age	−0.116	0.768***	−0.115	0.788***
	(0.112)	(0.195)	(0.118)	(0.195)
Urban x Age^2 / 100	0.248	−0.728***	0.247	−0.746***
	(0.189)	(0.182)	(0.200)	(0.182)
Region x Education	−0.031**	−0.133***	−0.033***	−0.132***
	(0.012)	(0.019)	(0.013)	(0.019)

Table A4 continued from previous page

	Child WFA Selection	Parent WFA Selection	Child Job Selection	Parent Job Selection
Region x Vocational	0.009	0.249***	−0.012	0.228***
	(0.042)	(0.076)	(0.044)	(0.077)
Region x Female	0.053	0.062	0.027	0.063
	(0.034)	(0.042)	(0.042)	(0.052)
Region x Age	−0.116**	0.104	−0.125**	0.107
	(0.049)	(0.066)	(0.057)	(0.068)
Region x Age^2 / 100	0.193**	−0.084	0.207**	−0.087
	(0.083)	(0.062)	(0.096)	(0.064)
Urban x Region x Education	0.055***	0.068***	0.039**	0.069***
	(0.017)	(0.026)	(0.018)	(0.027)
Urban x Region x Vocational	−0.141**	−0.236**	−0.132**	−0.228**
	(0.058)	(0.100)	(0.059)	(0.103)
Urban x Region x Female	−0.084*	−0.028	−0.077*	−0.017
	(0.046)	(0.068)	(0.047)	(0.072)
Urban x Region x Age	0.078	−0.330***	0.081	−0.337***
	(0.065)	(0.117)	(0.070)	(0.117)
Urban x Region x Age^2 / 100	−0.146	0.311***	−0.151	0.317***
	(0.110)	(0.109)	(0.117)	(0.110)
Year Fixed Effect: 2012	0.254***	0.269***	0.132***	0.278***
	(0.021)	(0.026)	(0.019)	(0.027)

Table A4 continued from previous page

	Child WFA Selection	Parent WFA Selection	Child Job Selection	Parent Job Selection
Year Fixed Effect: 2014	0.116***	0.256***	0.114***	0.259***
	(0.018)	(0.028)	(0.019)	(0.030)
Year Fixed Effect: 2016	0.325***	0.434***	0.330***	0.431***
	(0.018)	(0.027)	(0.021)	(0.031)
Year Fixed Effect: 2018	0.387***	0.634***	0.382***	0.634***
	(0.018)	(0.027)	(0.021)	(0.030)
Intercept	4.490***	12.497***	3.970**	12.915***
	(1.197)	(2.799)	(1.645)	(2.897)
N.	15 599	13 712	15 599	13 712
Adj. R2	0.227	0.291	0.220	0.291

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$