Heterogeneous Income Profiles and Life-Cycle Bias in Intergenerational Mobility Estimation

Martin Nybom
Jan Stuhler

DISCUSSION PAPER SERIES

IZA

Heterogeneous Income Profiles and Life-Cycle Bias in Intergenerational Mobility Estimation

Martin Nybom
SOFI, Stockholm University

Jan Stuhler
*University College London
and IZA*

ABSTRACT

Heterogeneous Income Profiles and Life-Cycle Bias in Intergenerational Mobility Estimation[*]

Research on intergenerational income mobility is based on current income since data on lifetime income are typically not available for two generations. However, using snapshots of income over shorter periods causes a so-called life-cycle bias if the snapshots cannot mimic lifetime outcomes. Using uniquely long series of Swedish income data, we show that current empirical strategies do not eliminate such bias. We focus on the widely adopted generalized errors-in-variables model and find that the remaining bias is substantial (20% of the true elasticity from left-side measurement error at the most relevant age range). IV estimates suffer from even stronger life-cycle effects and do not provide an upper bound. Inconsistencies stem from the interaction of two factors: heterogeneity in income profiles cannot be fully accounted for, and idiosyncratic deviations from average profiles correlate with individual characteristics and family background. We discuss implications of our findings for other literatures that depend on measurement of long-run income and income dynamics.

JEL Classification: J62, D3, D31

Keywords: intergenerational mobility, intergenerational income elasticity, life-cycle bias, non-classical measurement error, generalized errors-in-variables model, heterogeneous income profiles

Corresponding author:

Jan Stuhler
University College London
Drayton House Room G13
Gordon Street
London WC1H 0AX
United Kingdom
E-mail: j.stuhler@ucl.ac.uk

Introduction

Transmission of economic status within families is often measured by the intergenerational elasticity between parents' and children's lifetime income. A large and growing literature has estimated this parameter in order to analyze the extent of intergenerational mobility across countries, groups and time.[1] Unfortunately, the estimates in the early literature suffered greatly from measurement error in lifetime income, and successive improvements of the methodology led to large-scale corrections.[2]

While the early estimates were severely attenuated from approximation of lifetime values by noisy single-year income data for parents, Jenkins (1987) identifies systematic deviations of current from lifetime values over the life cycle as an additional source of inconsistency.[3] Haider and Solon (2006) and Grawe (2006) show that the latter is empirically of great importance. Various refined methods to eliminate such *life-cycle bias* have recently been presented and the generalized errors-in-variables model proposed by Haider and Solon has been widely adopted in the literature.

Our contributions are as follows. First, we show that such refined methods do not eliminate life-cycle bias in intergenerational elasticity estimates. Second, we use Swedish income tax data to quantify the importance of life-cycle effects in both ordinary least squares (OLS) and instrumental variable (IV) estimates. Our data contain nearly complete income histories of both fathers and sons, allowing us to derive a benchmark estimate for the intergenerational elasticity and thus to directly expose the life-cycle bias. Third, we discuss how current procedures can be modified to reduce life-cycle bias. Fourth, we discuss our results in the more general context of income dynamics over the life cycle. We conclude that (unobserved) income profile heterogeneity is substantial, can have stark consequences, and is harder to address than is commonly believed.

We specifically focus on the generalized errors-in-variables model, which suggests that intergenerational elasticities can be consistently estimated if lifetime income is approximated by current income at a certain age. We find that this procedure improves estimates but that the life-cycle bias is substantially larger than the generalized model predicts. The model disregards some of the heterogeneity in income profiles, and can therefore not eliminate life-cycle bias in intergenerational mobility or other applications. The remaining bias from left-side measurement error alone amounts to about 20 percent of the true intergenerational elasticity (0.21 vs. 0.27), even under favorable conditions.[4]

We also analyze two other methods to address measurement error in lifetime income: we illustrate why the consideration of differential income growth across subgroups will not yield consistent estimates, and show that IV estimates suffer from even greater life-cycle effects than OLS estimates. IV estimators do therefore not provide an upper bound of the true parameter, contrary to such arguments in the previous literature.

Our results are hence rather pessimistic. They imply that current methods to compensate for incomplete income data are less successful than commonly believed, casting doubts on the accuracy of mobility estimates as well as on the validity of comparisons across populations.

---

[1] See Solon (1999) for a comprehensive evaluation of the early empirical literature. Recent surveys include Björklund and Jäntti (2009) and Black and Devereux (2011).

[2] For example, the intergenerational elasticity of earnings for fathers and sons in the U.S. was estimated to be less than 0.2 among early studies (surveyed in Becker and Tomes, 1986), ranged between about 0.3 and 0.5 in the studies surveyed in Solon (1999), and is estimated to be around 0.6 or above in more recent studies like Mazumder (2005) and Gouskova et al. (2010).

[3] For example, lifetime income of highly educated individuals might be systematically understated by current income at young age if income growth increases with education.

[4] Assuming that central parameters of the generalized errors-in-variables model are perfectly observed, so that current income is measured at the exact proposed age. Adding right-side measurement error aggravates the life-cycle bias further if fathers' and sons' incomes are measured at similar ages.

However, having a benchmark elasticity allows us to describe the direction and magnitude of the bias at different stages in the life cycle, and to provide recommendations for researchers. We find that annual income at a late age provides a more reliable base for application of the generalized errors-in-variables model, that averaging over multiple income observations reduces life-cycle bias, and that the treatment of missing and zero income observations has important consequences.

Life-cycle bias stems from a more general mechanism involving the interaction of two factors: heterogeneity in income profiles cannot be fully accounted for, and unobserved idiosyncratic deviations from average profiles correlate with individual and family characteristics. This mechanism is of importance for other literatures that depend on measurement of long-run income and income dynamics. Major examples among these include studies on the returns to schooling, and the extensive literature that relates measures of stochastic income shocks to consumption or other outcomes. We discuss both in some detail. We further present evidence that unexplained dispersion in income growth is at least partially due to latent heterogeneity instead of persistent stochastic shocks.

The next section describes the general methodology and identifying assumptions employed in the early literature. We then examine methods based on more recent contributions: the generalized errors-in-variables model theoretically in section 2 and empirically in section 3, IV methods and consideration of income dynamics across subgroups in section 4. Section 5 reviews implications for other literatures, section 6 concludes.

1   The Intergenerational Mobility Literature

The prototypical regression model in intergenerational mobility research is

$$y^*_{s,i} = \beta y^*_{f,i} + \epsilon_i, \tag{1}$$

where $y^*_{s,i}$ denotes log lifetime income of the son in family i, $y^*_{f,i}$ log lifetime income of his father, ϵ_i is an error term that is orthogonal to $y^*_{f,i}$, and variables are expressed as deviations from their generational means.[5] The coefficient β is interpreted as the intergenerational income elasticity.

Equations akin to (1) appear in two distinctive forms in the literature. First, as a *statistical relationship* to measure the *outcome* of interest, i.e. the degree of intergenerational mobility. Second, as a *structural relationship* to study *causal mechanisms* of intergenerational transmission, derived from an economic model as in Becker and Tomes (1979). The statistical relationship is typically based on broad ex-post measures of long-run economic status such as lifetime income. The structural relationship instead relates to the ex-ante concept "permanent income", since expectations on long-run status determine individual behavior.[6] For simplicity, our analysis relates to the statistical relationship, but incomplete measurement of long-run status impedes identification of both types.

---

[5]We use the terms earnings and income interchangeably (since the issues that arise are similar), and examine fathers and sons since this has been the baseline case in the literature. A growing literature exists on intergenerational mobility in other family dimensions (e.g mothers, daughters or siblings) and in other income concepts (such as household income), for which our analysis is likewise relevant.

[6]For various reasons these concepts are not always clearly distinguished. First, simple economic models assign one time period to each generation, so that the concept of permanent and lifetime income coincide. Second, permanent income is difficult to measure. Empirical analysis of the structural relationship is still based on ex-post measures of (current) income, and is then often similar to the statistical relationship. Third, some of the empirical work in the literature has lately adopted the term "permanent income" even while focusing on the measurement of outcomes.

Approximation of Lifetime Income

As commonly available data sets do not contain complete income histories for two generations, a major challenge is how to approximate lifetime income.[7] Let y_i be some observed proxy for unobserved log lifetime income of an individual in family i, e.g. a single-year observation, an average of multiple annual income observations, or a more complex estimate based on such annual incomes. Observed values are related to true values by

$$y_{s,i} = y_{s,i}^* + u_{s,i},$$

where $y_{s,i}^*$ is the unobserved true log lifetime income of the son in family i and $u_{s,i}$ is measurement error. Similarly, for the father we observe

$$y_{f,i} = y_{f,i}^* + u_{f,i}.$$

The probability limit of the OLS estimator from a linear regression of y_s on y_f can be decomposed into

$$plim \ \hat{\beta}_{approx} = \frac{Cov(y_f, y_s)}{Var(y_f)} = \frac{\beta \ Var(y_f^*) + Cov(y_f^*, u_s) + Cov(y_s^*, u_f) + Cov(u_s, u_f)}{Var(y_f^*) + Var(u_f) + 2 \ Cov(y_f^*, u_f)}, \quad (2)$$

where we used eq. (1) to substitute for $y_{s,i}^*$ and applied the covariance restriction $Cov(y_{f,i}^*, \epsilon_i) = 0$. It follows that the estimator can be down- or upward biased and that the covariances between measurement errors and lifetime incomes impact on consistency. The empirical strategies employed in the literature in the last decades can be broadly categorized in terms of changes in identifying assumptions about these covariances.

First Two Waves of Studies

The first wave of studies, surveyed in Becker and Tomes (1986), neglected the problem of measurement error in lifetime status. Often just single-year income measures were used as proxies for lifetime income, thereby implicitly assuming that

$$Cov(y_f^*, u_s) = Cov(y_s^*, u_f) = Cov(u_s, u_f) = Cov(y_f^*, u_f) = 0,$$

and

$$Var(u_f) = 0.$$

Classical measurement error in lifetime income violates the latter assumption, so that estimates suffered from large attenuation bias. Estimates of the intergenerational elasticity were therefore too low. This problem was recognized in Atkinson (1980) and then frequently addressed in the second wave of studies (surveyed in Solon 1999). But the assumption remained that measurement errors are random noise, independent of each other and of true lifetime income. That life-cycle variation had to be accounted for was recognized, but it was generally assumed that including age controls in the regression equation would suffice. The assumptions were therefore

$$Cov(y_f^*, u_s) = Cov(y_s^*, u_f) = Cov(u_s, u_f) = Cov(y_f^*, u_f) = 0,$$

and

$$Var(u_f) \neq 0.$$

If these hold, then the probability limit in eq. (2) reduces to

---

[7] Note that the availability of better data would not generally solve the identification problem, since data sets cannot contain complete income histories for contemporary populations.

$$plim \ \hat{\beta}_{approx} = \beta \ \frac{Var(y_f^*)}{Var(y_f^*) + Var(u_f)}.$$

This is the classical errors-in-variables model; inconsistencies are limited to attenuation bias caused by measurement error in lifetime income of *fathers*. In contrast, measurement error in sons' lifetime income is not a source of inconsistency in this model.[8] Researchers typically used averages of multiple income observations for fathers to increase the signal-to-noise ratio, but gave less attention to the measurement of sons' income.

Recent Literature

Recently the focus has been on the existence of non-classical measurement error. An early theoretical discussion can be found in Jenkins (1987). Analyzing a simple model of income formation, he finds that usage of current incomes in eq. (1) will bias $\hat{\beta}$ as income growth over the life cycle varies across individuals. He concludes that the direction of this *life-cycle bias* is ambiguous, that it can be large, and that it will not necessarily be smaller if fathers' and sons' incomes are measured at the same age.

Haider and Solon (2006) demonstrate that life-cycle bias can explain the previously noted pattern that intergenerational elasticity estimates increase with the age of sampled sons.[9] They show that the association between current and lifetime income varies systematically over the life cycle, contrary to a classical errors-in-variables model with measurement error independent of true values. Böhlmark and Lindquist (2006) find strikingly similar patterns in a replication study with Swedish data.

Haider and Solon also note that controlling for the central tendency of income growth in the population by including age controls in eq. (1) will not suffice, as variation around the average growth rate will bias estimates. Vogel (2006) provides an illustration based on the insight that highly educated workers experience steeper-than-average income growth. Since available data tend to cover annual incomes of young sons and old fathers, lifetime incomes of highly educated sons (fathers) will be understated (overstated). It then follows from eq. (2) that $\hat{\beta}_{approx}$ is biased further downwards than implied by the classical errors-in-variables model.[10] Indeed, $\hat{\beta}_{approx}$ can be negative in extreme cases, as our data confirm. Various refined estimation procedures have been proposed to address such life-cycle bias. We proceed to examine the most popular one in detail.

2   The Generalized Errors-in-Variables Model

Haider and Solon (HS) formulate a generalized errors-in-variables (GEiV) model that incorporates variation in the association between annual and lifetime income over the life cycle. Their empirical analysis documents that this variation is substantial, with important implications for analyses based on short-term income measures and, in particular, the intergenerational mobility literature. HS argue that left-side measurement error can cause substantial life-cycle bias in intergenerational elasticity estimates, but that it is innocuous for consistency

---

[8]The parameter β is identified if the number of income observations for fathers is sufficiently large, if a consistent estimator for the attenuation factor can be derived, or if moment restrictions on the measurement errors can be justified, e.g. by inferring the distribution of the measurement errors from a different data set.

[9]For a summary, see Solon (1999). Age-dependency of elasticity estimates could also arise if the dispersion in transitory income and thus the attenuation bias vary over the life cycle. Such variation has been documented in Björklund (1993) for Sweden, but Grawe (2006) finds that the observed age-dependency can be better explained by the existence of life-cycle bias.

[10]If educational achievement is correlated within families, and if high education tends to correspond to high lifetime income, we have $Cov(y_f^*, u_s) < 0$, $Cov(y_s^*, u_f) < 0$, $Cov(u_s, u_f) < 0$ and $Cov(y_f^*, u_f) > 0$.

if lifetime income of sons are proxied by annual income at a certain age. The GEiV model has been widely adopted in the literature, and the procedure of measuring income around a certain age (around midlife) frequently applied.[11]

The underlying intuition of the model is that, for two individuals with different income trajectories, there will nevertheless exist an age t^* where the difference between their log annual income equals the difference between their log (annuitized) lifetime income. HS argue that the classical errors-in-variables model holds at t^*. HS first focus on left-side measurement error and assume that $y_{s,i}^*$ is unobserved and proxied by $y_{s,it}$, log annual income of sons at age t. The GEiV model is given by

$$y_{s,it} = \lambda_{s,t} y_{s,i}^* + u_{s,it}, \tag{3}$$

where $\lambda_{s,t}$ is allowed to vary by age and $u_{s,it}$ is orthogonal to $y_{s,i}^*$. Regressing $y_{s,it}$ on $y_{f,i}^*$ by OLS, and using eqs. (3) and (1) to substitute for $y_{s,it}$ and $y_{s,i}^*$, yields

$$plim\ \hat{\beta}_t = \frac{Cov(y_{s,t}, y_f^*)}{Var(y_f^*)} = \beta\lambda_{s,t} + \frac{Corr(y_f^*, u_{s,t})\,\sigma_{u_{s,t}}}{\sigma_{y_f^*}}. \tag{4}$$

HS assume that

$$Corr(y_f^*, u_{s,t}) = 0, \tag{5}$$

and conclude that left-side measurement error is innocuous for consistency if the sample is restricted to annual income of sons around an age t^* where $\lambda_{s,t}$ is close to one.

However, we argue that assumptions akin to (5) will typically not hold since idiosyncratic deviations from average income trajectories correlate with individual and family characteristics. This is in particular problematic in intergenerational mobility estimation.

Note that for more than two workers we will generally *not* find an age t^* where annual income is an undistorted approximation of lifetime income. Figure 1 illustrates this by plotting log income trajectories for workers 1, 2 (as in Figure 1 in HS) and an additional worker 3. The horizontal lines depict log annuitized lifetime income, and differences in workers' log lifetime income are given by the vertical distances between these lines. At age t_1^* the distance between the annual income trajectories equals the distance between the horizontal lines for worker 1 and 2, and at age t_2^* for worker 1 and 3. There exists no age where these distances are equal for all three workers at once.[12]

This illustrates that the coefficient $\lambda_{s,t}$ is merely a population parameter that reflects how differences in annual income and differences in lifetime income relate *on average in the population*. Individuals will nevertheless deviate from this average relationship, so that their annual income still over- or understates their lifetime income. For intergenerational mobility studies it is crucial that $\lambda_{s,t}$ contains no information on if such deviations ($u_{s,it}$ from eq. 3) correlate within families. The assumption that $u_{s,it}$ and $y_{f,i}^*$ are uncorrelated would mean that the relation between annual and lifetime income does not depend on family background. However, there are reasons to expect such dependency: parents can transmit abilities, or influence their offspring's educational and occupational choice, all of which could affect the shape of income profiles over the life cycle.

---

[11]Among others, in Gouskova et al. (2010) for the US; Björklund et al. (2006, 2009) for Sweden; Nilsen et al. (forthcoming) for Norway; Raaum et al. (2007) for Denmark, Finland, Norway, the UK and the US; Nicoletti and Ermisch (2007) for the UK; Piraino (2007) and Mocetti (2007) for Italy. More examples are covered in the surveys of Björklund and Jäntti (2009) and Black and Devereux (2011).

[12]This result does not depend on any peculiarities in the income growth process. Even for a simple linear formation of log annual income as analyzed in HS, the difference between log income y_{it} and log annuitized lifetime income depends on the income growth rate γ_i; it will have the same value for at most two distinctive realizations of γ_i; and will therefore systematically differ across individuals at any age t (proof in Appendix A1).

Similar arguments apply on an aggregate level. We thus expect different subgroups of the population to have systematically different relationships between annual and lifetime income. For example, different educational groups might experience different income growth over the life cycle, such that their annual incomes systematically over- or understate their lifetime income relative to the population.[13]

Technically, we examine the validity of assumption (5) by deriving the elements of $u_{s,it}$ for a given income formation model and analyzing its relation to $y_{f,i}^*$. It turns out that $u_{s,it}$ is correlated with the regressor $y_{f,i}^*$ even for the simple log-linear income formation model employed in HS (see Appendix A.2). The probability limit of $\hat{\beta}_t$ does therefore not generally equal $\lambda_{s,t}\beta$ since this correlation causes an omitted-variable bias. Corresponding biases arise in the case of right-side measurement error in which unobserved lifetime income of fathers is approximated by annual income (see Appendix A.3) and if approximations are made for both fathers and sons (Appendix A.4).

3 The GEiV Model: Empirical Evidence

We use Swedish panel data containing nearly life-long income histories to provide direct evidence on the life-cycle bias that remains after application of the GEiV model. The size of the bias depends on two factors. First, the complexity of income dispersion in the population.[14] Second, if the income dispersion is caused by heterogeneity or stochastic shocks. The former more than the latter would cause idiosyncratic deviations from average income profiles to be correlated within families.[15] Our findings will thus also be indicative about how complex the dispersion in income profiles is, and if its underlying causes are deterministic or stochastic. We will return to these issues in section 5.

3.1 Data Sources and Sample Selection

To the best of our knowledge, Swedish tax registry data offer the longest panel of income data, covering annual incomes across 48 years for a large and representative share of the population. Moreover, a multi-generational register matches children to parents, and census data provide information on schooling and other individual characteristics. All merged together, the data provide a unique possibility to examine life-cycle bias in intergenerational mobility estimation using actual income histories.

To select our sample, we apply a number of necessary restrictions. As we mainly aim to make a methodological point, we follow the majority of the literature and limit our sample to sons and their biological fathers. To these we merge income data for the years 1960-2007.[16] Since most other income measures are available only from 1968, we use total (pre-tax) income, which is the sum of an individual's labor (and labor-related) earnings, early-age pensions, and net income from business and capital realizations.

Our main analysis is based on sons born 1955-1957. Earlier cohorts could be used, but then we would observe fewer early-career incomes for fathers. Conversely, later cohorts are not included, since we want to follow the sons for as long as possible. Moreover, to avoid

---

[13]Such correlation between u_{it} and education would imply that the GEiV model can, for example, not be readily applied in the return to schooling literature. We provide evidence supporting this in a later section.

[14]For example, if individuals merely differ in linear income growth then differences in log lifetime income are well approximated by differences in log current income around midlife for the whole population and the GEiV model would perform relatively well.

[15]Simulation studies as in Stuhler (2010) can illustrate these arguments but are not informative about the size of the bias in applications, since the bias varies strongly with unknown characteristics of the income generating process.

[16]Income data for the year 1967 are missing in the registry.

large differences in the birth year of fathers, we exclude pairs where the father was older than 28 years at the son's birth. Young fathers and first-born sons are thus over-represented in our sample. Although this is a limitation, we expect any detected bias for this particular sample to understate the bias in the population.[17] On other sampling issues, we adopt the restrictions applied by HS and Böhlmark and Lindquist (2006).[18]

Our data come with a couple of drawbacks. To maximize the length of the income histories we use the measure total income, whereas e.g. HS use labor earnings. However, total income is a highly relevant measure of economic status, approximation of lifetime status gives rise to the same methodological challenges, and Böhlmark and Lindquist find that total income and earnings yield similar results for the intergenerational mobility of sons. Further, the use of tax-based data could raise concern about missing data in the low end of the distribution if individuals have no income to declare. The Swedish system however provides strong incentives to declare some taxable income since doing so is a requirement for eligibility to most social insurance programs. Hence, this concern most likely only applies to a very small share of the population.

Our data also have many advantages. First, they are almost entirely free from attrition. Second, they pertain to all jobs. Third, in contrast to many other studies, our data are not right-censored. Fourth, we use registry data, which is believed to suffer less from reporting errors than survey data. Fifth, and most important, we have annual data from 1960 to 2007, giving us nearly career-long series of income for both sons and their fathers. Overall, we believe that the data are the best available for the purpose of this study.

Our main sample consists of 3504 pairs of fathers and sons, with all sons' income measured from age 22 to age 50 and all fathers' income measured from age 33 to age 65, irrespective of birth years. Table 1 reports descriptive statistics. Rows (2) and (3) show that dispersions in lifetime income are of similar magnitudes for fathers and sons. Rows (4) and (5) provide information on the number of positive income observations. On average there are more than 28 observations for sons, and more than 30 for fathers, with relatively low dispersion in both cases.

3.2   Empirical Strategy

To assess the size of life-cycle bias, we compare estimates based on annual incomes with a benchmark estimate that is based on lifetime incomes. As in the theoretical discussion, we focus on left-side measurement error, although we provide brief evidence on life-cycle bias due to right-side and measurement error on both sides in a later subsection. We do this for two reasons. First, left-side measurement error has until recently been neglected in the literature. Second, life-cycle bias is not confounded by attenuation bias from classical measurement error on the left-hand side, which simplifies the analysis.

We first compute log lifetime incomes $y^*_{f,i}$ and $y^*_{s,i}$ using our series of income data. We use these to estimate eq. (1), yielding our benchmark estimate $\hat{\beta}$.[19] We then approximate

---

[17]Reduced sample heterogeneity will tend to decrease heterogeneity in income profiles, which in turn diminishes the idiosyncratic deviations from sample average relationships between annual and lifetime income that cause life-cycle bias.

[18]We restrict the sample to fathers and sons who report positive income in at least 10 years. We exclude those who died before age 50, and sons who immigrated to Sweden after age 16 or migrated from Sweden on a long-term basis (at least 10 years). Incomes are in 2005 prices, and an annual discount rate of 2 percent is used to calculate the discounted present value of lifetime income. Discounting procedures should adjust for economic growth if year of birth varies substantially. Otherwise β will partially reflect shared experience of economic growth through the correlation between father's and son's year of birth.

[19]Of course, this estimate is not exactly *true* since we still lack some years of income. This is however irrelevant for the validity of our approach to use it as benchmark. The GEiV model is not restricted to any specific population, and should therefore be applicable to our variant of the Swedish population in which we

log lifetime income of sons $y^*_{s,i}$ by log annual income $y_{s,it}$ (left-side measurement error) to reestimate eq. (1) separately for each age t, yielding a set of estimates $\hat{\beta}_t$. Finally, we estimate eq. (3), which provides us with estimates of $\lambda_{s,t}$.

Under the assumptions of the GEiV model, the probability limit of $\hat{\beta}_t$ equals $\lambda_{s,t}\beta$, and using annual income of sons at age t^* where $\lambda_{s,t} = 1$ consistently estimates β.[20] As discussed in the previous section, we anticipate $\hat{\beta}_t$ to be biased even after adjustment by $\hat{\lambda}_{s,t}$. The *remaining* life-cycle bias after such adjustment by the GEiV model, denoted by $b(\hat{t}) = \hat{\beta}_t/\hat{\lambda}_{s,t} - \hat{\beta}$, is thus of central interest.[21] Note that we assume that $\hat{\lambda}_{s,t}$ is known in order to evaluate the model's theoretical capability to adjust for life-cycle bias under favorable conditions. A second (known) source of inconsistency can arise in that the age profile of $\lambda_{s,t}$ will typically not be directly observed by the researcher.

3.3   Empirical Evidence

We first present estimates of $\lambda_{s,t}$. Figure 2 shows that $\hat{\lambda}_{s,t}$ rises over age and crosses one at around age $t^* = 33$. Largely consistent with others, we find that income differences at young (old) age substantially understate (overstate) differences in lifetime income.[22]

Our central estimates are presented in Figure 3, which plots $\hat{\beta}$ (the benchmark elasticity), $\hat{\beta}_t$ (estimates based on annual income of sons at age t), and $\hat{\beta}_t/\hat{\lambda}_{s,t}$ (estimates at age t adjusted by the GEiV model). The sample is balanced at each age, such that zero or missing income observations that are not considered for estimation of one coefficient are not used for estimation of the other coefficients. Hence the estimated benchmark elasticity varies slightly by age. We list our key findings.

First. Our benchmark estimate of the intergenerational elasticity of lifetime income for our Swedish cohort is about 0.27. This is marginally higher than what most previous studies have found for Sweden, and should be closer to the population parameter due to our nearly complete income profiles.[23]

Second. We confirm that the variation of $\hat{\beta}_t$ over age resembles the pattern in $\hat{\lambda}_{s,t}$, as predicted by HS. We therefore find that $\hat{\beta}_t$ increases with age and that the life-cycle bias is negative for young and positive for old ages of sons. One of the central predictions of the GEiV model, that current income around mid-life is a better proxy for lifetime income than income in very young or very old ages, is thus confirmed.

Third. The magnitude of life-cycle bias stemming from left-side measurement error alone can be striking. For example, analysis based on annual income of sons only two years below age t^* yields $\hat{\beta}_t = 0.191$, in contrast to a benchmark estimate that is almost 40 percent larger.

---

truncate income profiles at some age. It is nevertheless advantageous that we have long income histories. First, our benchmark estimate will be closer to the true value. Second, since the income profiles contain most of the idiosyncratic heterogeneity that leads to life-cycle bias, our estimate of the bias will be representative for a typical application. We provide evidence that our main findings are not sensitive to the exact length of observed income histories in section 3.4.

[20]Since age is a discrete variable, $\lambda_{s,t}$ will not necessarily equal exactly one at t^*. We adjust $\hat{\beta}_t$ according to eq. (3) by $\hat{\lambda}^{-1}_{s,t}$ at all ages, inlcuding t^*.

[21]The arguments of HS relate to the probability limit. In a finite sample we need to consider the distribution of $b(\hat{t})$. Reported standard errors for $b(\hat{t})$ are based on a Taylor approximation and take the covariance structure of $\hat{\beta}$, $\hat{\beta}_t$, and $\hat{\lambda}_{s,t}$ into account.

[22]Bhuller et al. (2011) find a very similar t^* for Norway. This is reassuring since they use labor earnings as income measure and observe somewhat longer income histories (ages 20-58) for a single generation. Our result that income differences at old age overstate differences in lifetime income differs from results for cohorts born in the 1930s in both HS and Böhlmark and Lindquist (2006), but are in line with results for cohorts born in the 1950s in Böhlmark and Lindquist, and Bhuller et al.

[23]Our benchmark elasticity is nevertheless still likely to understate the true intergenerational elasticity. We lack some early observations of fathers and late observations of sons, which reduces $\sigma_{y^*_s}$ and increases $\sigma_{y^*_f}$, thereby reducing the numerator and increasing the denominator of the OLS estimator.

Moreover, analysis based on income below age 26 yields a *negative* elasticity. We therefore find direct evidence on the importance of life-cycle bias in intergenerational mobility estimates that has been discussed in the recent literature.

Fourth. The life-cycle bias is larger than implied by the GEiV model. While the adjustment of estimates according to this model leads on average to sizable improvements, it cannot fully eliminate the bias. This holds true even under the assumption that the central parameters $\lambda_{s,t}$ are perfectly observed. The remaining bias is overall substantial, and especially large for young ages. Intergenerational elasticity estimates based on income at very young ages are still negative.

Fifth. The life-cycle bias is not minimized at age t^*, the age at which the current empirical literature aims to measure income, but at an age $t > t^*$. We report a similar pattern for other cohorts in section 3.4.

Sixth. The remaining life-cycle bias $\hat{b}(t)$ around age t^* is substantial and significantly different from zero. Table 2 shows that $\hat{b}(t)$ is on average around 0.05 over ages 31-35, which corresponds to about 20 percent of our benchmark. Furthermore, the large deviation from this average at age 32 indicates that the year-to-year variation can be large. Knowledge of age t^* will thus not eliminate life-cycle bias.

We briefly compare these empirical results with our theoretical discussion of the determinants of $\hat{b}(t)$. Table 3 decomposes $\hat{b}(t)$ according to eq. (4). Variation of $\hat{b}(t)$ over age stems mostly from variation in the residual correlation $Corr(y_f^*, u_{s,t})$, while the ratio $\sigma_{u_{s,t}}/\lambda_{s,t}\sigma_{y_f^*}$ is close to one over most of the life cycle.[24] Seemingly small residual correlations can thus translate into substantive biases. For example, a residual correlation of 0.03 translates into a life-cycle bias of more than 10 percent of the benchmark elasticity.

These results provide guidance for applied research, but some remarks about generalizability are warranted. Life-cycle bias will differ quantitatively across populations. The bias is determined by the degree of systematic differences in income profiles between sons from poor and sons from rich families. This mechanism is likely to vary across cohorts and countries. The question is if observed qualitative patterns over age can nevertheless be generalized. Figure 3 demonstrates that annual income at late age provides a more reliable base for application of the GEiV model in intergenerational studies than income at young age. The remaining life-cycle bias is large and negative up until the early forties, but then small for most older ages.[25] Thus, the relationship between current and lifetime income differs with respect to family background particularly at the beginning of the life cycle. This result is intuitive if one considers potential causal mechanisms of intergenerational transmission. Sons from rich families might acquire more education or face different conditions that particularly affect initial job search (e.g. regarding credit-constraints, family networks, or ex-ante information on labor market characteristics). Such mechanisms are likely to apply to some degree to most populations. Although the size of the life-cycle bias is bound to differ across populations, its pattern over age is thus likely to hold more generally. This conclusion is supported by results for other Swedish cohorts, as we will discuss later on.

3.4   Extensions

We proceed to examine alterations of the estimation procedure to reduce the bias, test the sensitivity of our results, and discuss if adjustments according to the GEiV model can eliminate life-cycle bias in other applications.

---

[24]The previously documented increase in $\lambda_{s,t}$ over age is offset by an increase in $\sigma_{u_{s,t}}$.

[25]The latter result cannot easily be exploited. Adjustment of $\hat{\beta}_t$ by $\hat{\lambda}_{s,t}$ can rarely be done in practice due to lack of information on the latter. Importing estimates of $\lambda_{s,t}$ from other sources can be misleading since its pattern over age could differ across populations.

Multi-Year Averages of Current Income

Some recent studies that reference to the GEiV model (see footnote 11) average over multiple income observations of sons, although without clear theoretical motivation. One rationale could be that researchers do not know the exact age at which $\lambda_{s,t}$ equals one. Our result that life-cycle bias is substantial even if this age would be known raises the question if and how such practice can help to reduce the bias.

We therefore estimate β_t using three-, five- and seven-year averages of son's income centered around age t^*. These averages are also used to estimate $\lambda_{s,t}$, and the remaining life-cycle bias after adjustment by $\hat{\lambda}_{s,t}$. The results are presented in Table 4. The remaining life-cycle bias falls in the number of income observations but is not eliminated. For the seven-year average, the estimated bias (in absolute value) is on average slightly below 0.03 at ages 31-35 compared to about 0.05 using one-year measures. The residual variance $\hat{\sigma}_{u_{s,t}}$ decreases by about a third when moving from one- to seven-year measures, and diminishes the estimated bias proportionally. The residual correlation falls only slightly and estimates of $\lambda_{s,t}$ remain stable. These improvements are moderate, but they still lead us to recommend averaging over multiple income observations on the left-hand side when possible.

Treatment of Outliers in the Income Data

Intergenerational elasticity estimates can be sensitive to how one treats outliers in general, and observations of zero or missing income in particular (Couch and Lillard, 1998; Dahl and DeLeire, 2008). We test the robustness of our results along this dimension by (i) balancing the sample across ages such that only sons with positive income in all ages 31-35 are included, (ii) bottom-coding very low non-missing incomes, and (iii) top-coding very high incomes.[26] We compare the life-cycle bias for ages 31-35 for each of these samples (presented in Table Table 5) with the results for our main sample in Table 2.

Estimates of the remaining life-cycle bias are on average about a third lower for the balanced sample than for our main sample (at ages 31-35), but still correspond to more than 10 percent of the benchmark elasticity. Decreases in both the residual correlation and residual variance contribute to this drop.[27] Bottom-coding has the opposite effect and increases the bias slightly since observations with zero income are now always included. Finally, results for a sample with top-coded incomes are very similar to those for the main sample, implying low sensitivity to the exact measurement of high incomes. While we thus find that zero and missing incomes are influential for the size of life-cycle bias, it is not obvious what the right sampling choice would be. To derive a general measure of mobility one would like to include all individuals, but our analysis shows that doing so comes with the cost of increased life-cycle bias.

Length of Observed Income Profiles

Although our data are to our knowledge the best available for our purpose, it might be a concern that our measures of lifetime income are still based on incomplete income histories. We thus perform a number of robustness tests. We consider a younger cohort — sons born

---

[26]As of the log-specification we do not expect high extremes to have as large influence as low extremes. Top-coding has however been suggested to test the sensitivity to some changes of administrative routines and tax levels across our time period (see Böhlmark and Lindquist, 2006).

[27]Excluding those with occasional zeros or missings reduces the number of extreme values and thereby the variation in $u_{s,it}$. The residual correlation decreases since individuals with frequent zero and missing income observations are likely to experience quite different income profiles than the average population, and therefore amplify the heterogeneity in income profiles that causes the residual correlation.

1958-60 — to study the influence of early-age income data of fathers, and an older cohort — born 1952-54 — to study the influence of late-age data of sons.

Age profiles of the life-cycle bias before and after adjustment by $\hat{\lambda}_{s,t}$ are shown in Figures 4 (main sample), 5 (cohort 1958-60), and 6 (cohort 1952-54) for variations of the age spans. Abstracting from general cohort differences, we find that changes in the fathers' age span have little effect on the life-cycle bias, probably due to our focus on left-side measurement error. In contrast, changes in the sons' age span cause noticeable shifts. This is not unexpected since changes in the age span on which our measures of lifetime income are based are likely to alter both $\sigma_{y_s^*}$ and $\lambda_{s,t}$ slightly. While the exact relation between the size of the life-cycle bias and age therefore depends on the definition of the age span, the major facts remain stable: the remaining life-cycle bias after adjustment by $\hat{\lambda}_{s,t}$ can be large and tends to be negative for young ages and around t^*.

Cohort and Population Differences

We use the same three cohort groups to briefly assess if the magnitude of life-cycle bias can be expected to vary across populations. To separate true cohort differences from differences due to age span definitions, we limit the income profiles of both fathers and sons to the longest age span observed in all three samples. We thus use incomes of sons for ages 22-47, and incomes of fathers for ages 36-65.[28] Differences between these samples are hence due to their respective data generating processes. Table 6 presents the most central results around age t^* for each sample.[29]

The 1958-60 cohort has an estimated benchmark elasticity $\hat{\beta}$ that is similar to our main cohort but a slightly larger remaining life-cycle bias $\hat{b(t)}$. For the 1952-54 cohort both $\hat{\beta}$ and $\hat{b(t)}$ are substantially lower. The differences in $\hat{b(t)}$ — substantial even for large samples and a fixed sampling procedure across cohorts within Sweden — confirm that life-cycle bias should be expected to differ across studies and populations also after adjustment by $\hat{\lambda}_{s,t}$.

The GEiV Model in Other Applications

The GEiV model can be applied to other literatures that use short-run income to proxy for unobserved long-run values. However, we argue that the model will not eliminate life-cycle bias in such other applications since assumptions akin to (5) on the covariance between residuals and the explanatory variable are generally unlikely to hold.

To provide brief evidence, we examine if the residuals from eq. (3) correlate with a range of characteristics that are of interest in various literatures, specifically (i) father's age at birth of his son, (ii) father's education, (iii) son's education, (iv) son's cognitive ability, and (v) son's country of birth. Table 7 describes how each variable is measured and presents the results. As expected, most of these correlations are significantly different from zero. Importantly, the residual correlations are non-zero also around age t^*. Knowing this age, or the age profile of $\lambda_{s,t}$, does therefore not allow researchers to fully control for life-cycle effects. The residuals correlate much stronger with son's and father's education than with father's log lifetime income, indicating that the GEiV model might perform worse in applications in which education plays a central role.[30] The correlations tend to be smaller when sons'

---

[28]Restricting the age intervals reduces the benchmark estimate. Dropping income observations for sons at old age and fathers at young age decreases $\sigma_{y_s^*}$ and increases $\sigma_{y_f^*}$, reducing the numerator and increasing the denominator of the OLS estimator.

[29]More detailed evidence on cohort differences are provided in Figures 11 and 12.

[30]Correlation with the schooling variables is large in particular for young age, reflecting that income growth of the highly educated is relatively strong while initial income is low.

income is measured at later ages, again supporting our argument that the GEiV performs better when applied to current income at ages $t > t^*$.

3.5 Measurement Error on the Right-Hand Side or Both Sides

Although our findings on left-side measurement error are conceptually interesting, evidence on the combined effects of life-cycle bias from both sides is more relevant for practitioners. The questions arise whether we find similar life-cycle effects from the right-hand side, and whether these tend to cancel out or aggravate the effects from left-side measurement error. Our data allow us to directly examine these questions. We now base estimates of β_t on lifetime income of sons and approximation of lifetime income by annual income for fathers (right-side measurement error) or approximation for both fathers and sons (measurement error on both sides). The probability limit of $\hat{\beta}_t$ is then affected by attenuation and life-cycle bias. We adjust for both according to the GEiV model. Results are shown in Figures 7 and 8.[31]

Figure 7 demonstrates the additional large attenuating effects from right-side measurement error. The remaining life-cycle bias after adjustment by the GEiV model follows a similar qualitative pattern over age as for the case of left-side measurement error. Figure 8 shows the remaining life-cycle bias in the case of measurement error on both sides with fathers' and sons' incomes measured at similar ages. It is overall larger than for left-side measurement error alone, thus indicating aggravating effects of measurement error on both sides.[32] Importantly, this is also the case when fathers' and sons' incomes are measured at their respective t^*. We again find that the GEiV model is less successful in eliminating the bias for early ages and around t^* than for later ages. Moreover, the estimates suffer from strong year-to-year variability, supporting the argument that comparisons of elasticity estimates across populations are likely to be difficult. Reducing this variability is an additional motive for averaging over multiple income observations on both sides, apart from our previous finding that it reduces the size of the bias.

4 Other Methods to Address Life-Cycle Bias

We briefly examine two other methods employed in the intergenerational mobility literature. We show that instrumental variable (IV) estimators do not yield an upper bound for β due to life-cycle effects, and discuss why consideration of differential income growth rates across subgroups also yields inconsistent estimates.

4.1 Instrumental Variable Methods

IV methods have been proposed as an alternative way to tackle attenuation bias that stems from right-side measurement error in eq. (1). Furthermore, in the form of two-sample IV

---

[31] Adjustment is based on separate estimates of λ_t for both fathers and sons, denoted $\hat{\lambda}_{f,t}$ and $\hat{\lambda}_{s,t}$. According to the GEiV model the probability limit of $\hat{\beta}_t$ equals $\theta_{f,t}\beta = (\lambda_{f,t}\sigma_{y_f^*}^2/(\lambda_{f,t}^2\sigma_{y_f^*}^2 + \sigma_{u_{f,t}}^2))\beta$ for right-side and $\lambda_{s,t}\theta_{f,t}\beta$ for both-side measurement error (assuming $Corr(u_{s,it}, u_{f,it}) = 0$ in addition to assumption 5). Therefore the remaining life-cycle biases equal $\hat{b(t)} = \hat{\beta}_t/\hat{\theta}_{f,t} - \hat{\beta}$ and $\hat{b(t)} = \hat{\beta}_t/\hat{\lambda}_{s,t}\hat{\theta}_{f,t} - \hat{\beta}$, respectively. See Appendix A.3 and A.4 for a detailed derivation of the components of these biases. For presentational purpose we use only one age subscript t and display combinations of annual income for sons and fathers with equal distances to their respective t^* in Figure 8

[32] This holds true if estimates are only adjusted for attenuation bias but not for life-cycle effects according to the GEiV model (see Figure 13). These results confirm and substantiate the theoretical predictions of Jenkins (1987) that measuring fathers' and sons' income at similar ages might not necessarily reduce life-cycle bias, and contradict arguments in the recent literature that such "life course matching" generally leads to smaller biases than asymmetric age combinations.

(TSIV) they are heavily relied on for countries with less rich data.[33] Little is however known about life-cycle effects in IV estimates.

Instruments used in the literature include indices of father's socioeconomic status (Zimmerman, 1992) and father's education (Solon, 1992). Following Solon (1992), assume that son's lifetime income is determined by

$$y^*_{s,i} = \eta_1 y^*_{f,i} + \eta_2 E_{f,i} + \varepsilon_i,$$

where father's education $E_{f,i}$ is allowed to have a direct relation with son's income in addition to the indirect relation via father's income. We are interested in the probability limit of the estimated coefficient $\hat{\beta}^{IV}_{approx}$ from a two-stage least-squares regression of $y_{s,i}$ (with $y_{s,i} = y^*_{s,i} + u_{s,i}$) on $y_{f,i}$ (with $y_{f,i} = y^*_{f,i} + u_{f,i}$).[34] In the classical errors-in-variables model, assuming that E_f is uncorrelated with u_f and u_s (assumption *IV1*), we have

$$
\begin{aligned}
plim\, \hat{\beta}^{IV}_{approx} &= \frac{Cov(E_f, y_s)}{Cov(E_f, y_f)} = \eta_1 + \eta_2 \frac{Var(E_f)}{Cov(E_f, y^*_f)} \\
&= \beta + \eta_2 \left(\frac{Var(E_f)}{Cov(E_f, y^*_f)} - \frac{Cov(E_f, y^*_f)}{Var(y^*_f)} \right) \\
&= \beta + \eta_2 \frac{\sigma_{E_f}}{Corr(E_f, y^*_f)\sigma_{y^*_f}} \left(1 - Corr(E_f, y^*_f)^2 \right),
\end{aligned}
$$

where the next to last step follows from the omitted-variable formula (see Solon, 1992, for a detailed derivation).[35] The IV estimator is then an upper bound for β if father's education has a non-negative relation with son's income (ass. *IV2*) and if its correlation with father's lifetime income is bounded between zero and one (ass. *IV3*).

However, we argue that assumption *IV1* that $E_{f,i}$ is uncorrelated with $u_{f,i}$ and $u_{s,i}$ does typically not hold, e.g. since education is correlated with income growth rates. Typical panel data cover income observations of young sons and old fathers. Then lifetime income of highly educated sons (fathers) will be understated (overstated), even if we control for the central tendency of income growth in the population. If education is correlated within families, then $Cov(E_f, u_f) > 0$ and $Cov(E_f, u_s) < 0$. Allowing for non-classical measurement error, the probability limit is more generally given by

$$
plim\, \hat{\beta}^{IV}_{approx} = \beta + \eta_2 \left(\frac{Var(E_f)}{Cov(E_f, y_f)} - \frac{Cov(E_f, y^*_f)}{Var(y^*_f)} \right) + \frac{Cov(E_f, u_s)}{Cov(E_f, y_f)} - \eta_1 \frac{Cov(E_f, u_f)}{Cov(E_f, y_f)}. \tag{6}
$$

The example above illustrates that the latter two terms can be negative, so $\hat{\beta}^{IV}_{approx}$ will not necessarily be an upper bound for β. The results from Grawe (2006) and Haider and Solon (2006) indicate that IV estimates might bound β if they are adjusted by a certain parameter (λ_t in the GEiV model), or if they are based on income at an age at which this parameter equals one. Our results imply that this is not the case since measurement errors in lifetime income are not orthogonal to education in either case (see section 3.4).

---

[33]TSIV was first applied to the intergenerational mobility literature by Björklund and Jäntti (1997).

[34]For generality we here use the notation from section 1 and exclude age subscripts.

[35]The relationship between the population parameter β in eq. (1) and η_1 and η_2 is given by

$$\beta = \frac{Cov(y^*_s, y^*_f)}{Var(y^*_f)} = \frac{Cov(\eta_1 y^*_f + \eta_2 E_f + \epsilon, y^*_f)}{Var(y^*_f)} = \eta_1 + \eta_2 \frac{Cov(E_f, y^*_f)}{Var(y^*_f)}$$

Empirical Evidence on Life-Cycle Bias in IV Estimates

We use our Swedish data to examine life-cycle bias in IV estimates. As most of the literature, we instrument for father's income by his years of education.[36] We again first derive a benchmark estimate $\hat{\beta}^{IV}$ using our measures of lifetime income y_f^* and y_s^*, to assess if IV methods provide an upper bound if lifetime incomes would be truly observed. The IV benchmark estimate $(\hat{\beta}^{IV} = 0.309, \hat{\sigma}_{\hat{\beta}^{IV}} = 0.037)$ is larger than our OLS benchmark estimate $(\hat{\beta} = 0.273, \hat{\sigma}_{\hat{\beta}} = 0.017)$, indicating that assumptions *IV2* and *IV3* can indeed be expected to hold. However, the difference $\hat{\beta}^{IV} - \hat{\beta}$ is statistically insignificant.

To probe the validity of assumption *IV1*, we again focus on left-side measurement error (using son's annual income at age t) in order to abstract from attenuation bias and to directly compare life-cycle bias in OLS and IV estimates. Figure 9 plots these estimates together with the OLS benchmark $\hat{\beta}$ and reveals two important results.

First, life-cycle effects from left-side measurement error are substantially larger in IV than in OLS estimates.[37] Adjustment by λ_t would thus improve IV estimates only modestly. Usage of education as an instrument aggravates the life-cycle bias since income profiles differ strongly with education — the covariances in eq. (6) are thus relatively large (c.f. section 3.4). Second, IV estimates are well below the benchmark also at t^* at age 33 $(\hat{\beta}_{t^*}^{IV} = 0.183, \hat{\sigma}_{\hat{\beta}_{t^*}^{IV}} = 0.056$ whereas $\hat{\beta} = 0.270, \hat{\sigma}_{\hat{\beta}} = 0.017)$. We therefore conclude that *absent life-cycle effects*, IV estimates bound the true parameter from above. But since applications are typically based on current income, IV estimates do not bound β in practice. Given the large sensitivity of IV estimates to the age at which sons' incomes are measured (ranging between 0.08 and 0.53 over ages 30-45), we argue that comparisons of IV estimates across populations are not robust if based on short spans of income data.

4.2 Life-Cycle Patterns in Income Across Subgroups

An alternative method to address life-cycle bias is to model life-cycle income processes across subgroups. For example, Vogel (2006) models income growth across different skill groups instead of assuming a uniform growth rate, and argues that the resulting measures of lifetime income eradicates life-cycle bias in intergenerational elasticity estimates.

Education is an important determinant of individual income growth, so the procedure can reduce measurement error in lifetime income that arises from idiosyncratic deviations from average income growth. However, after accounting for differential income growth across educational groups, other determinants of income will lead to deviations from the mean income growth rate within the given group. Since such determinants might often be shared by members of the same family, measurement errors in income growth rates are still likely to be correlated within families. For example, if a father holds an occupation that typically leads to steeper than average income growth, then his son's income growth might also be steeper because he is relatively more likely to enter the same occupation.[38]

This example indicates that estimates could be improved by considering additional individual characteristics for the estimation of growth rates in more specific subgroups. But the crucial insight is that we will not be able to sufficiently project life-cycle trajectories of

---

[36]We impute years of education from data on level of educational attainment as recorded in the 1970 census, i.e. when the fathers were around 40 years old. Using level dummies yields similar results.

[37]In contrast, life-cycle effects from right-side measurement error are not particularly strong in IV estimates (figure available from the authors upon request). According to eq. (6) this is partly attributable to the downscaling effect of η_1 $(\hat{\eta}_1 = 0.263$ for our sample).

[38]Within-family correlation in the choice of profession can be observed, for example, in the likelihood of becoming president of the United States. More comprehensive evidence based on a somewhat larger sample is given in Corak and Piraino (2010).

income, since individual growth rates are determined by both observable and unobservable characteristics that can correlate within families. Unexplained dispersion in income growth is large: for example, Jenkins (2009) finds substantial deviations of individual income trajectories from average trajectories of groups defined by education, sex, and birth cohort in British data. The remaining life-cycle bias caused by within-family correlation of the unexplained part of individual income growth rates should therefore not be expected to be negligible.

Our data allow us to provide evidence in support of this argument. We derive average growth in log income for various groups of sons by regressing current log income on a polynomial in age. Figure 10 depicts such income trajectories for four groups of sons defined by education (non-college/college) and their father's lifetime income (below/above median). We find that while income trajectories are simply shifted for the two groups without college education, the difference in life-cycle patterns is substantial for the other two groups: college-educated sons with richer fathers have much stronger income growth than college-educated sons with poorer fathers. We thus find evidence for deterministic heterogeneity in income profiles even after controlling for a range of observable characteristics (sex, cohort, age, country of birth and education).

This result is of interest beyond the intergenerational mobility literature, in particular for the ongoing debate of whether idiosyncratic differences in income profiles of otherwise observationally equivalent individuals are mainly due to deterministic heterogeneity or persistent stochastic shocks (see e.g. Guvenen, 2007 for a recent contribution). These two sources are generally hard to distinguish from one another. We argue that the intergenerational dimension provides a novel perspective on this debate: systematic life-cycle patterns in income that are transmitted within families cannot stem from unexpected stochastic shocks, but have to be caused by deterministic factors. We thus conclude that unobserved deterministic heterogeneity in income profiles should be expected to be large, at least among highly educated individuals.[39] We expand on this issue in the next section.

5   Income Dynamics in Other Fields

Life-cycle bias in intergenerational mobility estimation arises from the inability to fully address heterogeneity in income profiles. It follows that we need to understand individual income dynamics over the life cycle before income dynamics over generations can be understood. The intergenerational mobility literature thus faces similar difficulties as many other literatures that depend on measures of long-run income values and income dynamics. In this section we discuss our results in this more general context. In particular we comment on the applicability of the GEiV model in the returns to schooling literature and on the literature on income dynamics and stochastic shocks.

5.1   Income Dynamics and Stochastic Shocks

Permanent income[40] and income risk are thought to be essential determinants of individual behaviour.[41] Measures of these are commonly derived by statistical decomposition of observed

---

[39]This finding is consistent with Guvenen (2009), who finds that heterogeneous income profiles explain a larger fraction of income inequality among college-educated individuals.

[40]A reminder of terminology is useful to distinguish ex-ante and ex-post concepts of long-run income status. Friedman gave an extensive exposition on that the term "permanent income" was to denote a theoretical concept, an individual's *ex-ante* expectation on some *unspecified* long-run income measure that individual behaviour (consumption) is based on, not an *ex-post* realisation of a *specific* measure like lifetime income (Friedman 1957, p.23). In the intergenerational mobility literature this distinction is not always clearly made, see footnote 6.

[41]A large literature explores the link between innovations in income and consumption. Income risk has increasingly also been related to other outcomes, including wealth inequality, saving, asset prices, crime,

income profiles into a part explained by observable individual-specific covariates and an auto-correlated error term.[42] The error term is often split into a random walk ("permanent component" plus "permanent shock") and an $AR(q)$ process ("transitory component").[43] Such decompositions equate unexplained dispersion in income with stochastic uncertainty: individuals with stronger persistent income growth relative to others with similar observable characteristics are thought to have experienced stronger positive permanent income shocks.

But what the econometrician, who has relatively little information on individual characteristics, deems as unexpected shock might not be a shock to the individual. An alternative strand of the literature suggests that individual income profiles instead differ deterministically as of unobserved heterogeneity, e.g. due to individual ability.[44] Such latent heterogeneity cannot be equated with uncertainty — for example, Cunha et al. (2005) argue that a substantial part of the variability in ex-post returns to schooling is predictable by individuals. Income shock measures could thus be subject to measurement error that stems from confusion of deterministic heterogeneity and stochastic uncertainty. The discussion of measurement issues and income profile heterogeneity in the intergenerational mobility literature becomes interesting in this context.

First, an intergenerational perspective allows us to distinguish deterministic income heterogeneity from stochastic persistent shocks, since only the former will be (partially) transmitted within families.[45] In section 4.2, we showed that family background can explain a substantial share of unexplained dispersion in income profiles even after controlling for a wide range of observable characteristics. This result favors the view that unobserved heterogeneity in income profiles can be quantitatively important.

Second, measurement error in income shocks should be expected to be non-classical, since unobserved determinants of income growth (e.g. cognitive or non-cognitive abilities) will correlate with shock measures. The impact of such measurement error is not restricted to attenuation bias (as is often assumed in the literature) and can enter observed relationships between outcomes and measures of income shocks in more complicated ways.[46]

Third, note that these problems arise even when the information set of the individual and the econometrician coincide. Cunha et al. (2005) refer to the variability in the returns to schooling that is unpredictable to the individual as reflecting 'luck'. We argue that it can also reflect deterministic returns to characteristics that the individual is not aware of. For example, abilities might be known or unknown to the individual but can in either case affect life-cycle income growth.

---

health, and child outcomes.

[42]An alternative strategy, presented in Blundell and Preston (1998), is to model the link between income and consumption in order to infer permanent and transitory shocks from consumption data.

[43]Variations and extensions of such decomposition procedure include MaCurdy (1982), Abowd and Card (1989), Moffitt and Gottschalk (1995), and Meghir and Pistaferri (2004).

[44]This strand of the literature, originating in Lillard and Weiss (1979), allows for heterogeneity in trends. Baker (1997) provides evidence that profile heterogeneity models provide a more consistent representation of earnings dynamics. Guvenen (2007; 2009) argue that accounting for heterogeneous income profiles yields more accurate predictions about life-cycle consumption behavior. Browning et al. (2010) present a model with heterogeneity in many additional parameters.

[45]Of course, we might expect that persistent stochastic shocks in parents' income affect their offspring. But such transmission within families would be deterministic, and would thus not represent (unexpected) stochastic shocks for the offspring. Likewise, individuals might tend to experience similar income shocks within families (for example if family members tend to work in the same industry or region), but if their income profiles show similarities over many years (for example because both parents and offspring work in a shrinking industry) then these income changes are unlikely to be unexpected to the offspring.

[46]The consequences depend on the examined structural relationship. For example, it is not obvious why ability should have an autonomous impact on consumption besides its relationship through income growth. But correlation of permanent income shock measures with ability seems problematic for many other outcomes that have recently been related to income shocks.

Finally, the intergenerational mobility literature exemplifies well how stark the consequences of non-classical measurement error can be and how difficult it is to address such measurement error. Revisions of estimation strategies in this literature were largely data driven. It is a reason for caution that such data-driven validation is more difficult in the literature that relates stochastic income shocks to various outcomes — availability of longer income histories is informative about ex-post lifetime measures, but not necessarily about individuals' expectations on long-run income status.

Of course, inference does often not depend on variation in individual income shocks, but on variation on an aggregated scale (such as cohorts or regions). While the exact underlying assumptions differ across applications, our view is that it can be summarized as an instrumental variable problem: researchers search for dimensions (cohorts, regions, etc.) that cause variation in the size of shocks, but relate to outcomes only via these shocks. Such aggregation will not necessarily eliminate measurement errors. For example, variation in income processes over regions might stem from pure "market luck" that is orthogonal to conditions within the region, but also from differences in the distributions of, or returns to, unobserved determinants of income. The latter would cause measures of income shocks to correlate with income determinants also on an aggregate scale.

5.2   Returns to Schooling

Empirical studies on the returns to schooling are potentially subject to life-cycle bias, as differences in current earnings can misrepresent lifetime returns. The issue is less pressing than for intergenerational mobility studies (the latter require data for two generations and researchers are thus more likely to observe non-representative age spans), but it becomes important under certain conditions. The actual relationship between current earnings and schooling over the life cycle might not follow the specified functional relationship.[47] Even with a flexible functional form one might fail to recover the true relationship if the age distribution of sampled individuals does not represent the age span of interest.

This becomes particularly relevant in cohort-based estimation. Heckman et al. (2006) argue that, given non-stationary labor markets, cohort analysis is preferred to an approach in which various cohorts are sampled (the so-called synthetic cohort approach) to estimate ex-post returns. Repeated cross-section or panel data are then required, in which researchers rarely observe complete lifetime earnings. The question is if life-cycle bias can be eliminated in estimates based on shorter earnings spans.

The GEiV model illustrates the difficulty of such reliance on short-term measures, but implies that earnings differences at a certain age approximates differences in lifetime earnings and thus lifetime returns to schooling well. However, our analysis shows that the model provides only a partial correction for life-cycle effects. Knowledge of the central parameter λ_t is insufficient, since the shapes of earnings profiles of different educational subgroups deviate from the population average grasped by λ_t. The projection error from eq. (3) thus correlates with education, which we confirm empirically in section 3.4. Bhuller et al. (2011) present a general analysis of life-cycle bias in returns to schooling estimates, and also analyze the applicability of the GEiV model in this context.

6   Conclusions

In this paper, we have demonstrated that recent methodological advances in the estimation of intergenerational income elasticities do not eliminate life-cycle bias. Current empirical

---

[47]For example, Heckman et al. (2006) argue that log earnings in schooling and experience are not multiplicative separable.

practice based on the generalized errors-in-variables model improves estimates, but we find that the remaining life-cycle bias is nevertheless substantial. The bias is strongly negative when using annual income below age 30, and remains negative in the mid-thirties. Since this is the age at which studies that rely on the generalized model have aimed to measure income, we expect resulting estimates to (still) understate the intergenerational elasticity.

One might argue that remaining life-cycle biases do not pose major problems, since interest often lies in the relative intergenerational mobility across populations (e.g. in cross-country studies). Estimates might be sufficiently precise if we believe that the magnitude of the remaining life-cycle bias is similar across countries. However, such belief would correspond to the assumption that within-family correlation in the shape of income profiles is of the same magnitude across countries. It is therefore a direct restriction on a potential characteristic of intergenerational mobility. Since countries differ with respect to the degree of intergenerational mobility, we also suspect that they differ in this specific aspect of it.

Comparisons of intergenerational elasticity estimates based on short-run income data across countries, groups or cohorts are thus of limited reliability, unless differences in estimates are large. Our finding that the life-cycle bias varies even across Swedish cohorts born in the same decade supports this conclusion. Moreover, comparisons are even less robust if based on IV estimates, which we find to suffer from larger life-cycle effects than OLS estimates.

Our analysis provides some guidance for applied research. We find evidence that annual incomes at later ages (e.g. age 40-50) provide a more reliable base for application of the GEiV model. The remaining life-cycle bias can be further reduced by averaging over multiple income observations for *both* fathers and sons, and the treatment of zero and missing income observations has important consequences. To derive a general measure of mobility one would like to include such observations, but doing so comes with the cost of increased vulnerability to life-cycle bias.

Further refinements of empirical practice with restricted use of income observations around a specific age can thus improve upon previous estimates, but will not eliminate life-cycle bias. Development of a more structured approach that aims to capitalize on all available income data seems thus desirable. Future research could in particular benefit from a more comprehensive exploitation of partially observed income growth patterns. Intergenerational mobility estimates are often based on multiple income observations per individual, but researchers typically disregard the idiosyncratic income growth across these observations. Such partially observed growth patterns are determined by all observable and unobservable characteristics of the individual. Hence they contain more information on the unobserved part of income profiles and thus on lifetime income than what current income levels and observable characteristics can provide.

Our results add to a general conclusion that can be drawn from the intergenerational mobility literature: addressing heterogeneity in income profiles is an important, difficult and recurrently underestimated task. The central problem is that idiosyncratic deviations from average income profiles correlate with a wide range of individual and family characteristics. We hope that our discussion underlines the potential severity of this issue for other literatures that rely on measurement of long-run income or income dynamics.

7   Tables and Figures

Table 1: Descriptive Statistics: Means and Standard Deviations by Son's Birth Year

	All	1955	1956	1957
Father's age at birth of son	24.68 (2.53)	24.66 (2.51)	24.77 (2.50)	24.62 (2.58)
log lifetime income (sons)	11.97 (0.43)	11.98 (0.42)	11.98 (0.42)	11.95 (0.44)
log lifetime income (fathers)	11.72 (0.42)	11.73 (0.44)	11.72 (0.43)	11.72 (0.40)
# of pos. income obs. (sons)	28.52 (1.86)	28.57 (1.71)	28.56 (1.74)	28.43 (2.11)
# of pos. income obs. (fathers)	30.32 (3.76)	29.99 (4.13)	30.36 (3.62)	30.59 (3.48)
Father-son pairs (N)	3504	1167	1173	1164

Notes: Lifetime income is annuitized by dividing total lifetime income by number of non-missing income observations.

Table 2: OLS Estimates of Elasticities and Life-Cycle Bias

t=Age	$\hat{\lambda}_{s,t}$	$\hat{\beta}$	$\hat{\beta}_t$	$\hat{\beta}_t/\hat{\lambda}_{s,t}$	$\hat{b(t)}$	$\hat{b(t)}$ in %	N
31	0.897	0.266	0.191	0.213	-0.053	19.8	3478
	(0.019)	(0.016)	(0.023)	(0.029)	(0.021)		
32	0.909	0.267	0.246	0.271	0.003	1.3	3476
	(0.018)	(0.016)	(0.023)	(0.028)	(0.021)		
33	0.982	0.267	0.203	0.207	-0.061	22.7	3479
	(0.020)	(0.016)	(0.025)	(0.031)	(0.024)		
34	1.039	0.256	0.212	0.204	-0.051	20.1	3469
	(0.019)	(0.016)	(0.025)	(0.031)	(0.023)		
35	1.114	0.261	0.234	0.210	-0.052	19.7	3460
	(0.021)	(0.016)	(0.027)	(0.029)	(0.022)		

Notes: Cohort group 1955-1957, left-side measurement error only. Standard errors in parentheses, which for $\hat{\beta}_t/\hat{\lambda}_{s,t}$ and $\hat{b(t)}$ are based on Taylor approximations that take the covariance structure of $\hat{\lambda}_{s,t}$, $\hat{\beta}$, and $\hat{\beta}_t$ into account. Column (7) displays $\hat{b(t)}$ in percent of our benchmark estimate $\hat{\beta}$.

Table 3: Decomposition of Life-Cycle Bias

t=Age	$\hat{b(t)}$	$Corr(y_f^*, \hat{u}_{s,t})$	$\hat{\sigma}_{u_{s,t}}$	$\hat{\sigma}_{y_f^*}$	$\hat{\sigma}_{u_{s,t}}/\hat{\lambda}_{s,t}\hat{\sigma}_{y_f^*}$
31	-0.053	-0.044	0.455	0.424	1.198
32	0.003	0.003	0.431	0.423	1.123
33	-0.061	-0.052	0.485	0.422	1.169
34	-0.051	-0.050	0.452	0.422	1.031
35	-0.052	-0.049	0.494	0.422	1.050

Notes: Cohort group 1955-1957, left-side measurement error only.

Table 4: OLS Estimates with Multi-Year Averages of Son's Income

t=Age	$\hat{\lambda}_{s,t_t}$	$\hat{\beta}$	$\hat{\beta}_t$	$\hat{\beta}_t/\lambda_{s,t}$	$\hat{b(t)}$	$\hat{\sigma}_{u_{s,it}}$	$Corr(\hat{u}_{s,it}, y^*_{f,i})$	N
				9A: Three-Year Average				
31	0.863	0.268	0.218	0.253	-0.015	0.339	-0.016	3496
	(0.013)	(0.016)	(0.020)	(0.026)	(0.017)			
32	0.948	0.267	0.214	0.226	-0.041	0.365	-0.045	3499
	(0.015)	(0.016)	(0.021)	(0.027)	(0.017)			
33	1.011	0.267	0.229	0.227	-0.041	0.372	-0.047	3495
	(0.015)	(0.016)	(0.022)	(0.030)	(0.022)			
34	1.078	0.268	0.229	0.212	-0.056	0.396	-0.064	3491
	(0.016)	(0.016)	(0.024)	(0.028)	(0.020)			
35	1.142	0.262	0.232	0.203	-0.059	0.411	-0.069	3486
	(0.017)	(0.016)	(0.025)	(0.028)	(0.020)			
				9B: Five-Year Average				
31	0.847	0.268	0.213	0.251	-0.017	0.290	-0.021	3500
	(0.012)	(0.016)	(0.018)	(0.025)	(0.014)			
32	0.925	0.268	0.229	0.248	-0.020	0.310	-0.025	3501
	(0.012)	(0.016)	(0.020)	(0.024)	(0.014)			
33	1.010	0.267	0.226	0.224	-0.043	0.336	-0.055	3500
	(0.013)	(0.016)	(0.021)	(0.026)	(0.016)			
34	1.057	0.267	0.235	0.222	-0.044	0.332	-0.060	3497
	(0.013)	(0.016)	(0.022)	(0.028)	(0.018)			
35	1.141	0.268	0.247	0.216	-0.052	0.384	-0.065	3495
	(0.015)	(0.016)	(0.024)	(0.028)	(0.019)			
				9C: Seven-Year Average				
31	0.866	0.270	0.221	0.255	-0.015	0.279	-0.020	3503
	(0.011)	(0.016)	(0.018)	(0.024)	(0.013)			
32	0.908	0.268	0.227	0.250	-0.018	0.261	-0.026	3501
	(0.010)	(0.016)	(0.018)	(0.023)	(0.013)			
33	0.975	0.268	0.238	0.245	-0.023	0.279	-0.034	3502
	(0.011)	(0.016)	(0.020)	(0.024)	(0.013)			
34	1.069	0.268	0.245	0.229	-0.038	0.318	-0.055	3502
	(0.013)	(0.016)	(0.022)	(0.025)	(0.015)			
35	1.122	0.267	0.251	0.224	-0.044	0.341	-0.061	3500
	(0.014)	(0.016)	(0.023)	(0.027)	(0.017)			

Notes: Cohort 1955-57, left-side measurement error only. Current income of sons are measured as three-, five-, and seven-year averages centered around age t. $\hat{\beta}_t$ is the coefficient from regressing the log of the average on father's log lifetime income. $\lambda_{s,t}$ is from the regression of the log of the average on the son's log lifetime income. Standard errors are in parentheses.

Table 5: Robustness Tests with Sample Variations

t=Age	$\hat{\lambda}_{s,t}$	$\hat{\beta}$	$\hat{\beta}_t$	$\hat{\beta}_t/\hat{\lambda}_{s,t}$	$\hat{b(t)}$	$\hat{\sigma}_{u_{s,it}}$	$Corr(\hat{u}_{s,it}, y^*_{f,i})$	N
				8A: Balanced Sample				
31	0.826	0.257	0.184	0.223	-0.033	0.408	-0.029	3415
	(0.018)	(0.015)	(0.021)	(0.028)	(0.021)			
32	0.839	0.257	0.227	0.271	0.014	0.398	0.012	3415
	(0.018)	(0.015)	(0.020)	(0.028)	(0.021)			
33	0.909	0.257	0.185	0.203	-0.053	0.445	-0.046	3415
	(0.020)	(0.015)	(0.023)	(0.032)	(0.025)			
34	0.964	0.257	0.219	0.227	-0.029	0.402	-0.030	3415
	(0.018)	(0.015)	(0.022)	(0.028)	(0.021)			
35	1.043	0.257	0.239	0.229	-0.027	0.457	-0.026	3415
	(0.020)	(0.015)	(0.024)	(0.027)	(0.020)			
				8B: Bottom-Coded Incomes				
31	0.913	0.271	0.205	0.225	-0.046	0.442	-0.040	3498
	(0.018)	(0.016)	(0.023)	(0.028)	(0.020)			
32	0.947	0.271	0.245	0.259	-0.012	0.434	-0.011	3500
	(0.018)	(0.016)	(0.023)	(0.029)	(0.020)			
33	0.999	0.270	0.201	0.201	-0.069	0.434	-0.067	3501
	(0.018)	(0.016)	(0.024)	(0.030)	(0.022)			
34	1.094	0.270	0.247	0.226	-0.044	0.437	-0.046	3501
	(0.018)	(0.016)	(0.025)	(0.029)	(0.020)			
35	1.153	0.270	0.250	0.217	-0.053	0.450	-0.057	3498
	(0.018)	(0.016)	(0.026)	(0.028)	(0.019)			
				8C: Top-Coded Incomes				
31	0.951	0.251	0.191	0.201	-0.050	0.448	-0.045	3478
	(0.019)	(0.016)	(0.024)	(0.027)	(0.020)			
32	0.959	0.252	0.246	0.257	0.005	0.426	0.004	3476
	(0.018)	(0.016)	(0.023)	(0.027)	(0.019)			
33	1.038	0.252	0.203	0.196	-0.056	0.477	-0.051	3479
	(0.020)	(0.016)	(0.025)	(0.029)	(0.023)			
34	1.104	0.240	0.213	0.193	-0.047	0.443	-0.050	3469
	(0.019)	(0.015)	(0.025)	(0.029)	(0.022)			
35	1.185	0.246	0.234	0.198	-0.048	0.484	-0.050	3460
	(0.021)	(0.015)	(0.027)	(0.027)	(0.020)			

Notes: Cohort 1955-57, left-side measurement error only. Panel A is balanced across ages, hence excluding all observations of individuals who have zero or missing incomes at any age 31-35. In panel B, annual incomes of sons below 10 000 SEK are bottom-coded up to 10 000 SEK (including zeros, but excluding missings). In panel C, annual incomes of sons above 2 million SEK are top-coded down to 2 million SEK. Standard errors are in parentheses.

Table 6: Summary of Cohort Differences, Averages over Ages 31-35

Cohort Group	$\hat{\lambda}_{s,t}$	$\hat{\beta}$	$\hat{\beta}_t$	$\hat{\beta}_t/\hat{\lambda}_{s,t}$	$\hat{b(t)}$	$\hat{b(t)}$ in %	N
1958-60	1.071	0.274	0.235	0.220	-0.054	19.9	3427
	(0.022)	(0.016)	(0.028)	(0.032)	(0.026)		
1955-57	1.066	0.246	0.216	0.204	-0.042	17.2	3444
	(0.020)	(0.015)	(0.024)	(0.028)	(0.020)		
1952-54	1.059	0.206	0.190	0.179	-0.027	12.8	3160
	(0.021)	(0.015)	(0.024)	(0.027)	(0.019)		

Notes: Left-side measurement error only. Table displays averages of estimates and standard errors (in parentheses) across ages 31-35. $\hat{b(t)}$ is significantly different from zero (p-value<0.05) at three ages (out of five) for 1958-60, at four ages for 1955-57, and at two ages for 1952-54. For all cohort groups, lifetime income is restricted to be measured over identical ages: 22-47 for sons, and 36-65 for fathers. Column (7) displays $\hat{b(t)}$ in percent of our benchmark estimate $\hat{\beta}$ (as average over the age interval).

Table 7: Correlations Between Residuals and Characteristics

	Age Interval of Sons				
	26-30	31-35	36-40	41-45	46-50
Father's log lifetime income	-0.057*	-0.050*	-0.063*	-0.020	-0.007
Father's age at birth of son	-0.054*	0.014	0.045*	0.017	-0.006
Father's education	-0.158*	-0.061*	-0.045*	0.035	0.028
Son's education	-0.278*	-0.112*	-0.002	0.085*	0.088*
Son's cognitive ability	-0.108*	-0.073*	-0.050*	0.022	-0.004
Son's country of birth	-0.040*	-0.026	-0.002	-0.032	0.028

Cohort 1955-1957. Table depicts correlations between characteristics listed in the first column and the residuals (as average in each five-year year age interval) from eq. (3). The education variables are years of education measured at about age 35, "Son's country of birth" is an indicator for being born outside Sweden, and "Son's cognitive ability" is a standardized cognitive ability measure from the military enlistment cognitive test at age 18. Superscript star indicates whether correlation has p-value<0.05.

Figure 1: Illustrative Example of Log Annual Income Trajectories
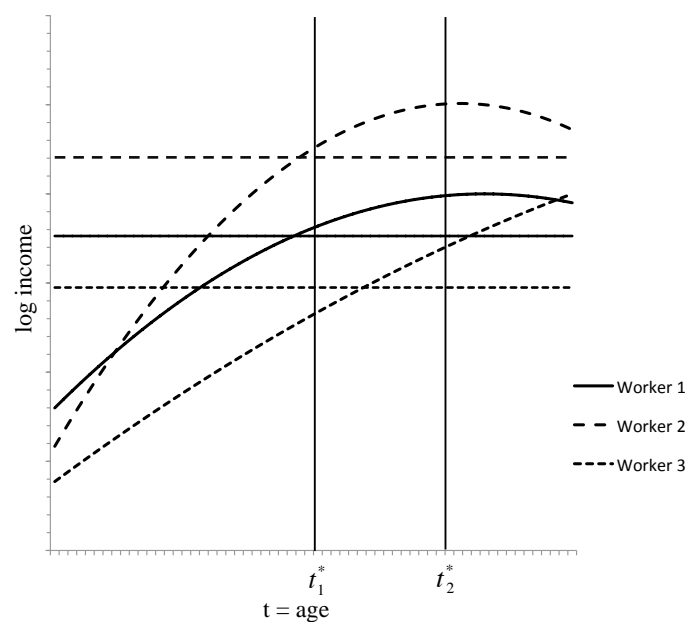
Notes: Illustrative Example. For each worker, the upward-sloping line depicts log annual income by age, the horizontal line depicts log annuitized lifetime income.
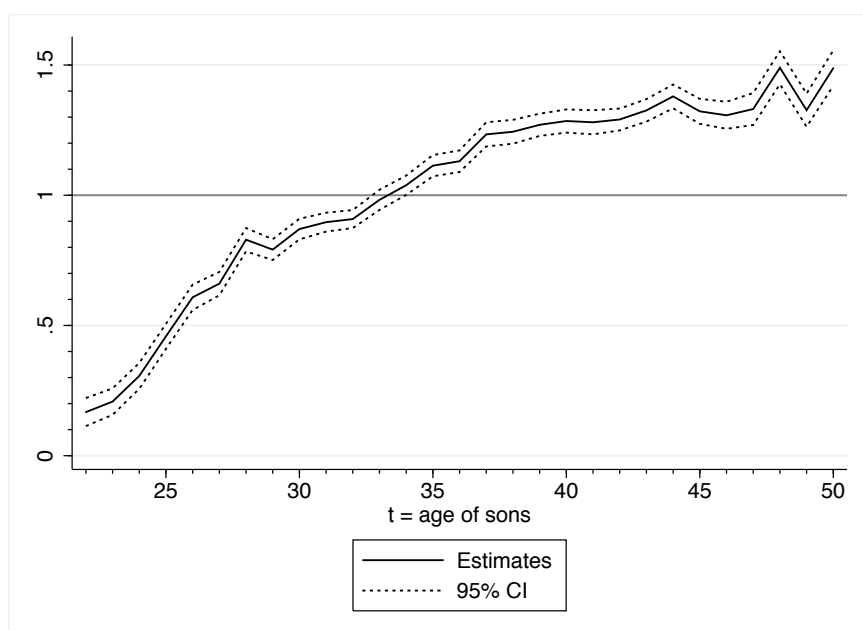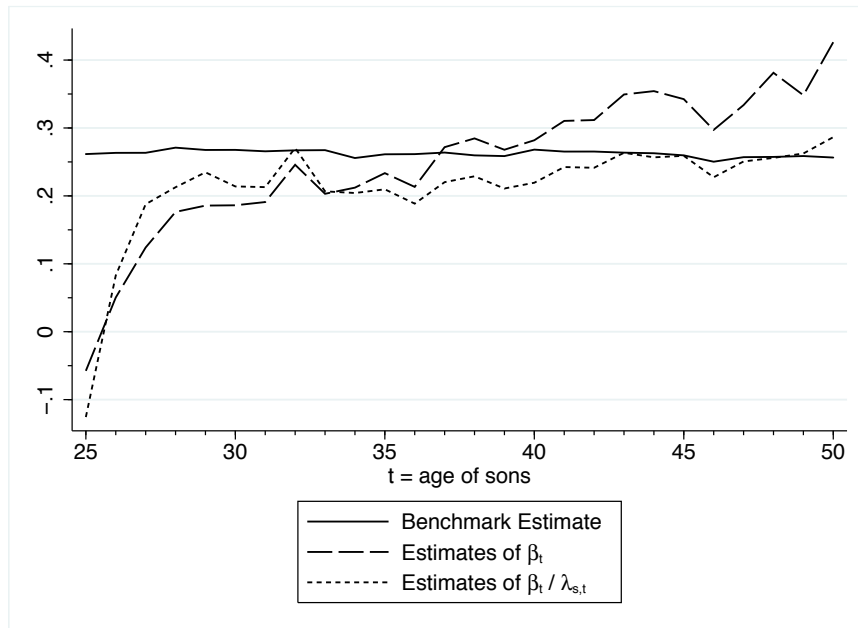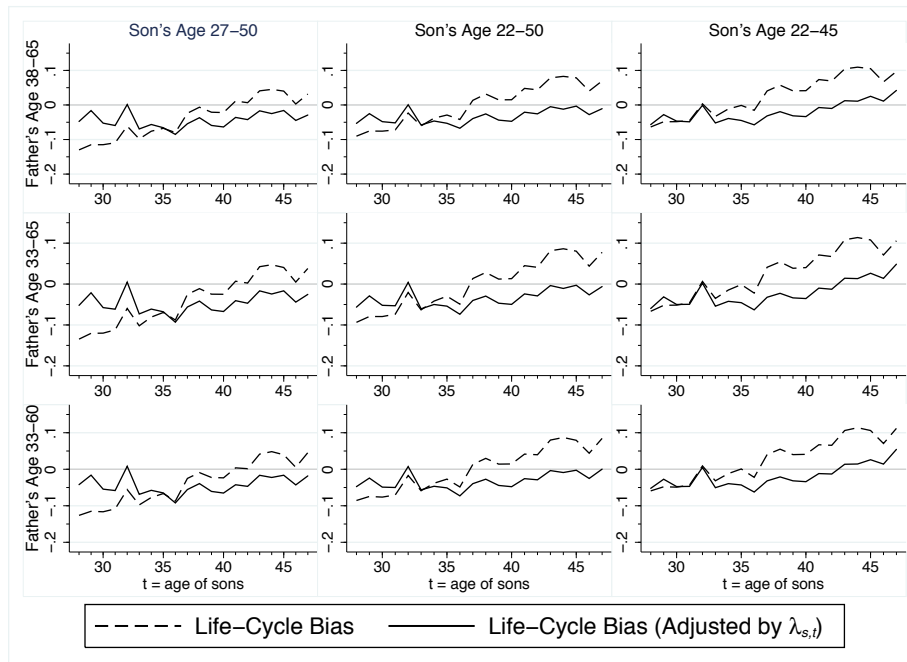
Figure 2: OLS Estimates of $\lambda_{s,t}$
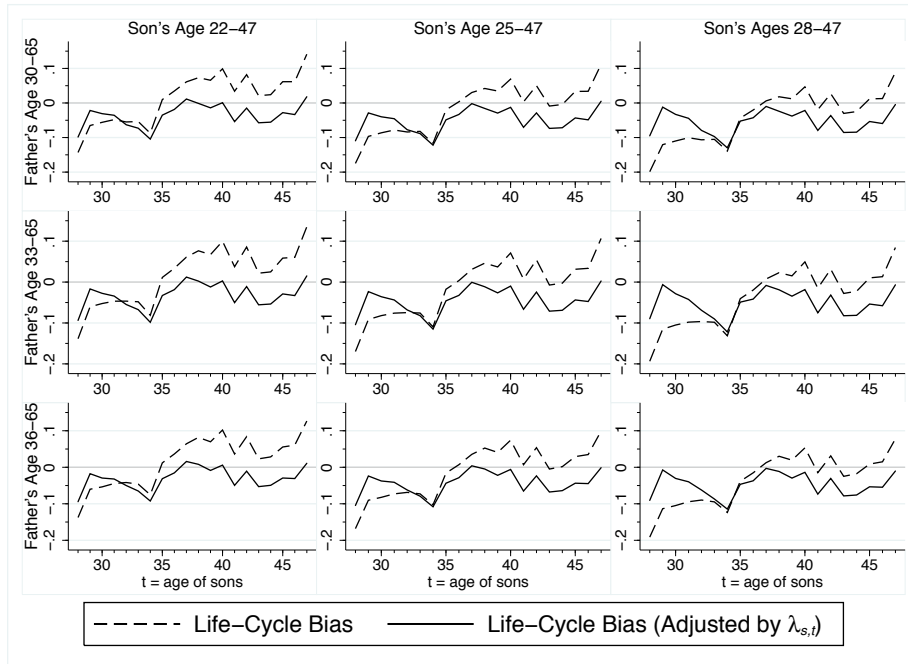
Figure 3: OLS Estimates of Elasticities and Life-Cycle Bias

Notes: Cohort 1955-57, left-side measurement error only.

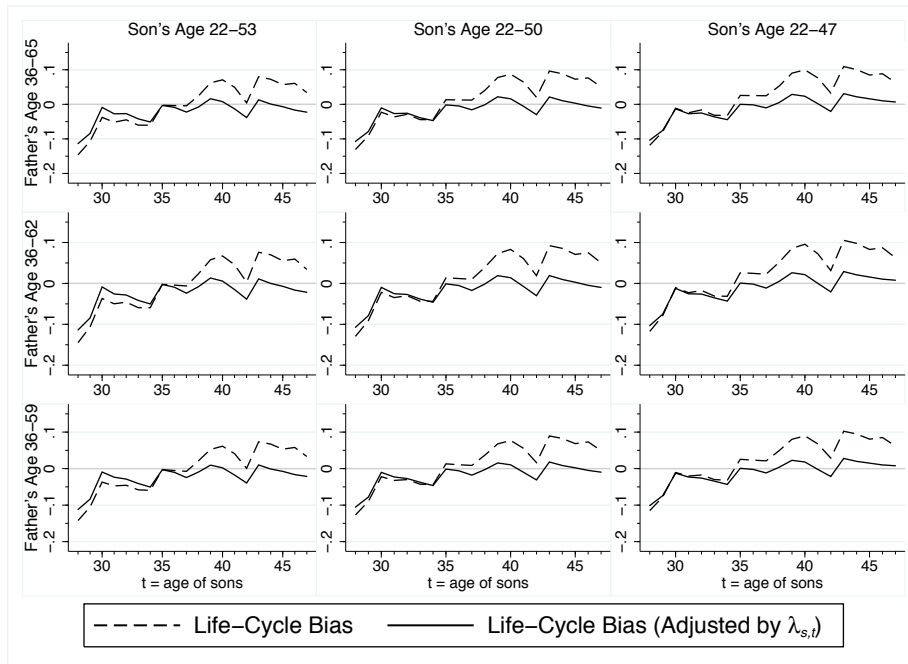Figure 4: Estimates of Life-Cycle Bias for Different Age Spans (Cohort 1955-57)


Notes: Left-side measurement error only. The age span of observed incomes of sons (fathers) varies along the horizontal (vertical) dimension.

Figure 5: Estimates of Life-Cycle Bias for Different Age Spans (Cohort 1958-60)

Notes: Left-side measurement error only. The age span of observed incomes of sons (fathers) varies along the horizontal (vertical) dimension.

Figure 6: Estimates of Life-Cycle Bias for Different Age Spans (Cohort 1952-54)

Notes: Left-side measurement error only. The age span of observed incomes of sons (fathers) varies along the horizontal (vertical) dimension.

Figure 7: OLS Estimates of Elasticities with Right-Side Measurement Error
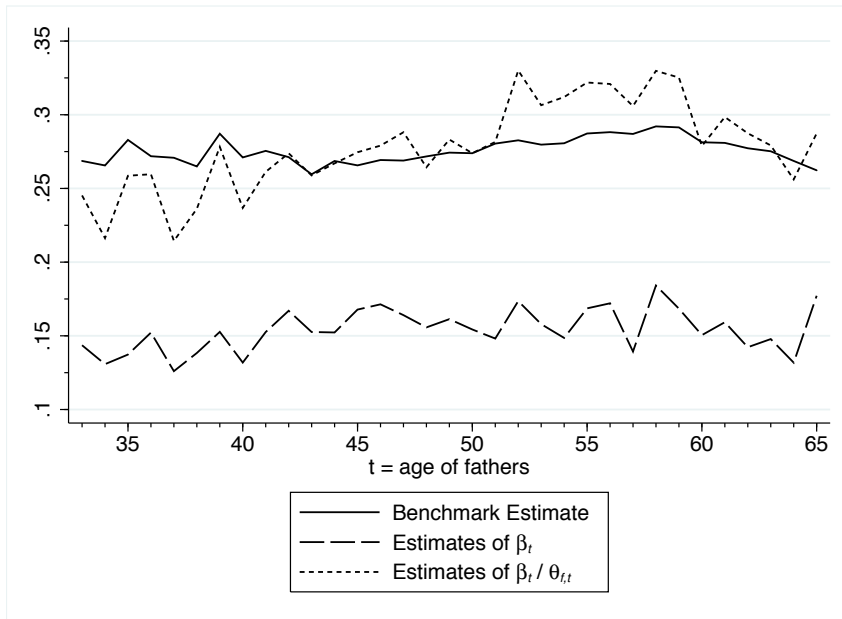
Notes: Cohort 1955-57, right-side measurement error only.

Figure 8: OLS Estimates of Elasticities with Both-Side Measurement Error
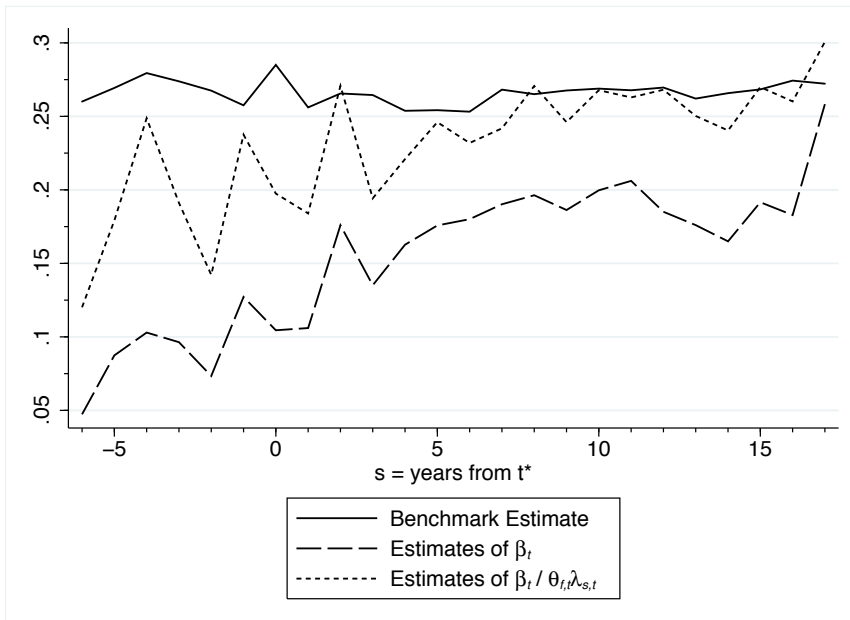
Notes: Cohort 1955-57, measurement error on both sides. To keep the analysis in two dimensions, we only display results for annual incomes at the same distance from t* for sons and fathers. At s=0 both are measured at their respective t*, at s=5 both are measured five years after t*, etc.

Figure 9: IV Estimates Compared with OLS and Benchmark
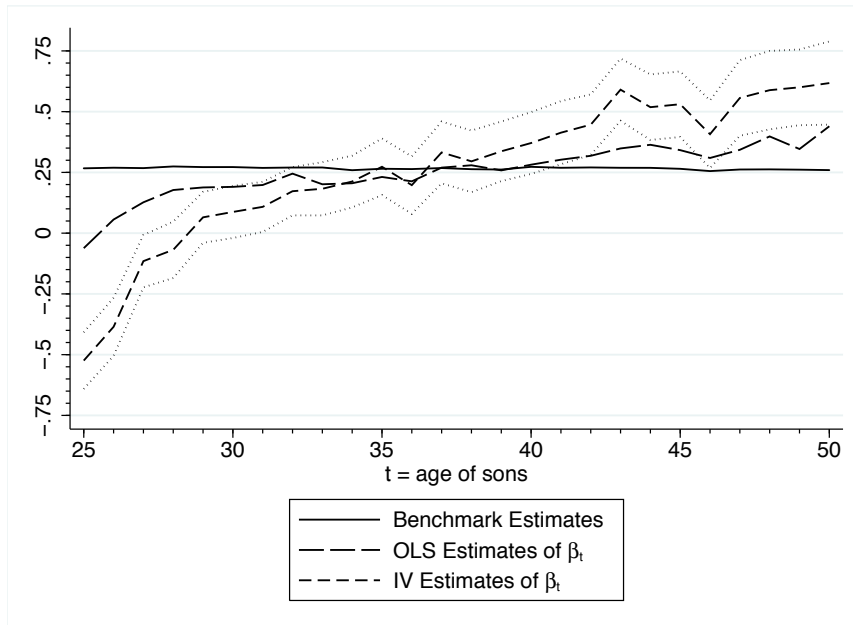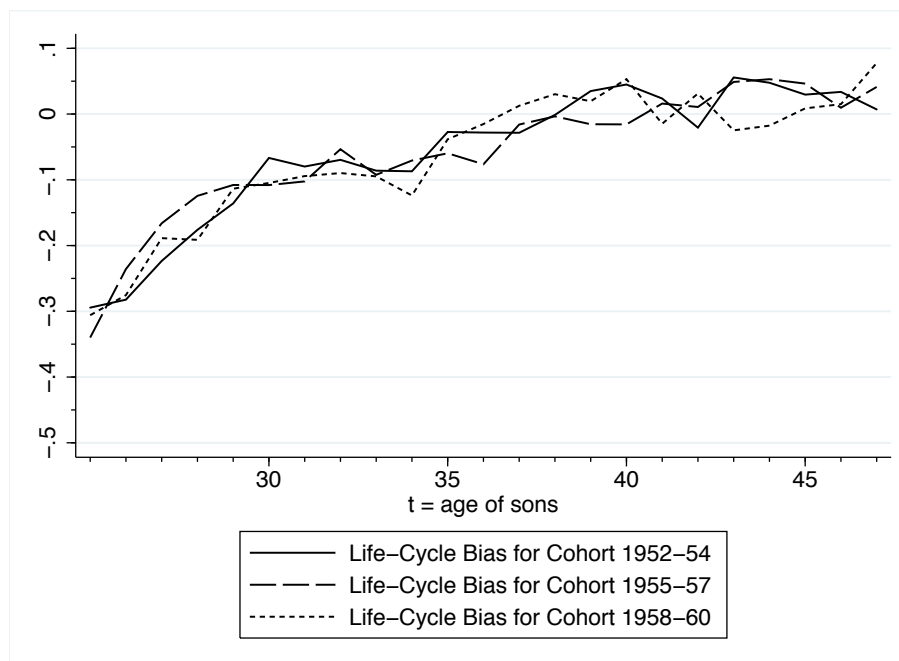

Notes: Cohort 1955-57, left-side measurement error only.

Figure 10: Life-Cycle Patterns in Income Across Subgroups
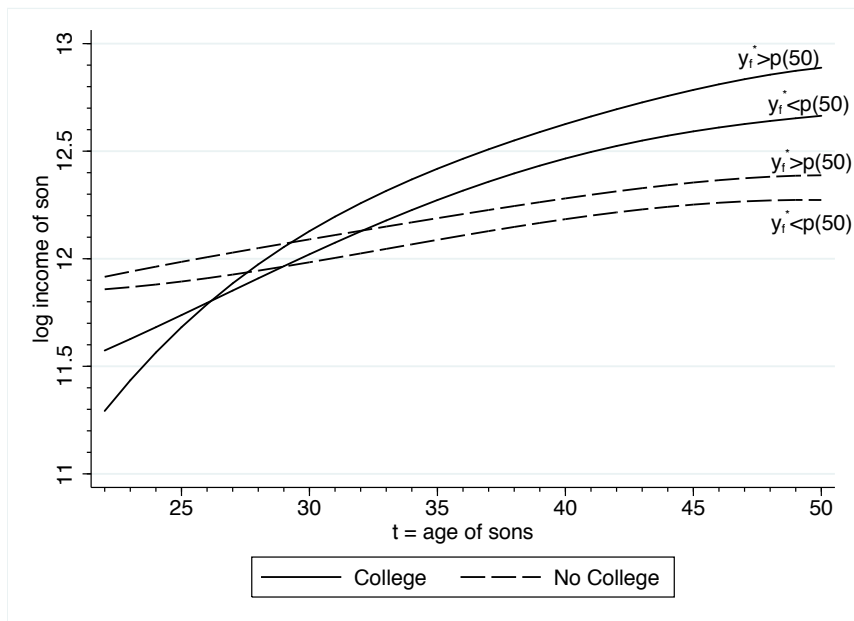
Notes: The trajectories depict average growth in log income over the life cycle for sons born in 1955-57, separately for sons with fathers above and below median lifetime income.

Figure 11: Cohort Differences in OLS Estimates of Life-Cycle Bias

Notes: Left-side measurement error only. For all cohorts, lifetime income is restricted to be measured over identical ages: 22-47 for sons, and 36-65 for fathers.

Figure 12: Cohort Differences in OLS Estimates of Remaining Life-Cycle Bias
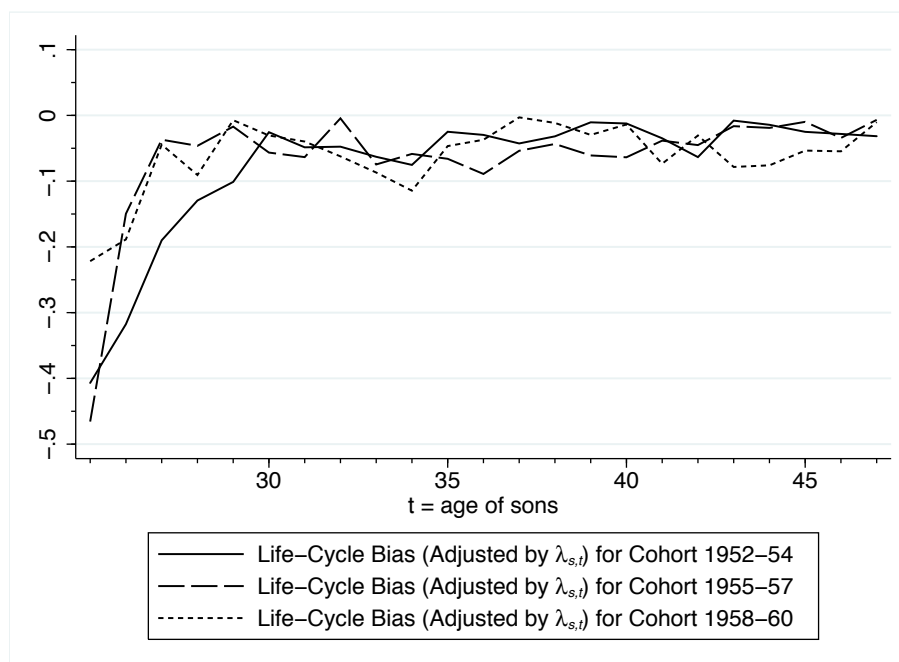
Notes: Left-side measurement error only. For all cohorts, lifetime income is restricted to be measured over identical ages: 22-47 for sons, and 36-65 for fathers.

Figure 13: Life-cyle Bias with Both-Side Measurement Error
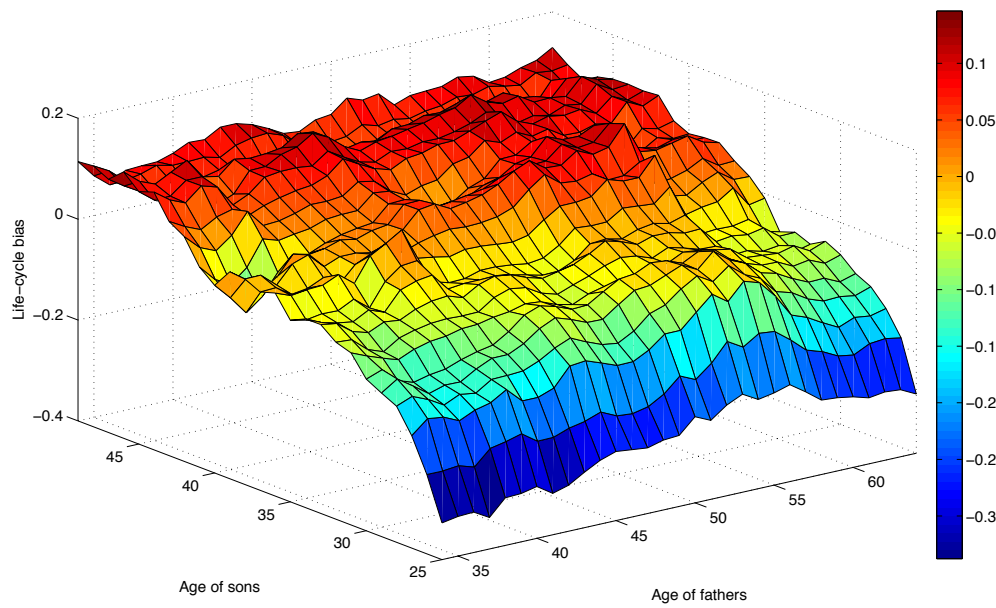
Notes: Cohort 1955-57, measurement error on both sides. Estimates are based on 3-year averages of annual income and are only corrected for attenuation bias (no application of the GEiV model).

8   Appendix

A.1 Annual and Lifetime Values Over the Life Cycle

Proposition. *(i) For all age t, the difference between log annual income y_{it} and the log of the annuitized value of the present discounted value of lifetime income varies with respect to the individual's income growth rate γ_i. (ii) For any given age t, the difference will be equal for at most two different realizations of γ_i.*

Proof. As in Haider and Solon (2006), suppose that log annual income of worker i at age t is given by

$$y_{it} = \eta_i + \gamma_i t \qquad (7)$$

For simplicity assume infinite lifetimes and a constant real interest rate $r > \gamma_i$. The annuitized value of the present discounted value of lifetime income, denoted B_i, is then

$$\sum_{s=0}^{\infty} exp(\eta_i + \gamma_i s)(1+r)^{-s} = \sum_{s=0}^{\infty} B_i (1+r)^{-s} = \frac{1+r}{r} B_i$$

Hence the log of the annuitized value equals

$$
\begin{aligned}
log\, B_i &= log\left(\frac{r}{1+r} \sum_{s=0}^{\infty} exp(\eta_i + \gamma_i s)(1+r)^{-s} \right) \\
&\cong log\, r + \eta_i - log(r - \gamma_i)
\end{aligned}
$$

The difference D_{it} between log annual income y_{it} and the log of the annuitized value of the present discounted value of lifetime income $log\, B_i$ is thus

$$D_{it} = \gamma_i t - log\, r + log(r - \gamma_i)$$

Depending on t, D_{it} decreases or increases in individuals' income growth rates γ_i,

$$\frac{\partial D_{it}}{\partial \gamma_i} = t - \frac{1}{r - \gamma_i}$$

The second derivative with respect to γ_i is negative,

$$\frac{\partial^2 D_{it}}{\partial^2 \gamma_i} = -(r - \gamma_i)^{-2} < 0$$

D_{it} is therefore a strictly concave function of γ_i conditional on t given, and a specific value of D_{it} can stem from at most two different values of γ_i. $\qquad\square$

A.2 Life-Cycle Bias: Left-Side Measurement Error

Assume that we wish to estimate the regression model (1), but that log lifetime income of sons $y_{s,i}^*$ is approximated by $y_{s,it}$, log annual income at age t. Fathers' log lifetime income $y_{f,i}^*$ is observed. We express the linear projection of $y_{s,it}$ on $y_{s,i}^*$ as in eq. (3). As shown in section 2, the probability limit of the OLS estimator of a linear regression of $y_{s,it}$ on $y_{f,i}^*$ is then

$$plim\, \hat{\beta}_t = \beta\, \lambda_{s,t} + \frac{Cov(y_f^*, u_{s,t})}{Var(y_f^*)}$$

Using the linear log income formation model given in eq. (7) we have

$$
\begin{aligned}
y_i^* &= log\left(\sum_{s=0}^{\infty} exp(\eta_i + \gamma_i s)^s (1+r)^{-s}\right) \\
&= \eta_i + log(1+r) - log(1+r - exp(\gamma_i))
\end{aligned}
\tag{8}
$$

and rewrite

$$
\begin{aligned}
u_{s,it} &= y_{s,it} - \lambda_{s,t} y_{s,i}^* \\
&= \eta_i + \gamma_i^s t - \lambda_{s,t}\left(\eta_i + log(1+r) - log(1+r - exp(\gamma_i))\right)
\end{aligned}
$$

where γ_i^f, γ_i^s denote income growth rates of fathers and sons. For simplicity assume that $\eta_i = \bar{\eta}$, and that $Var(\gamma_i^f) = Var(\gamma_i^s) = \sigma_\gamma^2$. The probability limit of $\hat{\beta}_t$ at age t^* for which $\lambda_{s,t} = 1$ is then[48]

$$
plim\, \hat{\beta}_t = \beta - \frac{Cov\left(log(1+r - exp(\gamma_i^f)), \gamma_i^s t^* + log(1+r - exp(\gamma_i^s))\right)}{Var\left(log(1+r - exp(\gamma_i^f))\right)}
$$

If lifetime income correlates within families we should also expect correlation of income growth rates within families. The second term in the previous expression is thus non-zero for a general class of joint distribution functions for γ_i^f, γ_i^s and represents the life-cycle bias. For example, assume that income growth rates are joint normally distributed

$$
\gamma = \begin{pmatrix} \gamma_i^f \\ \gamma_i^s \end{pmatrix} \overset{d}{\sim} N\left(\begin{pmatrix} \mu_\gamma \\ \mu_\gamma \end{pmatrix}, \begin{pmatrix} \sigma_\gamma^2 & \sigma_\gamma^2 \rho \\ \sigma_\gamma^2 \rho & \sigma_\gamma^2 \end{pmatrix}\right)
$$

and that r is large relative to μ_γ and σ_γ so that $r > \gamma_i$ still holds. Application of the Delta method on

$$
G(\gamma_f, \gamma_s) = \begin{pmatrix} g_1(\gamma^f, \gamma^s) \\ g_2(\gamma^f, \gamma^s) \end{pmatrix} = \begin{pmatrix} log(1+r - exp(\gamma_i^f)) \\ \gamma_i^s t^* + log(1+r - exp(\gamma_i^s)) \end{pmatrix}
$$

with

$$
\frac{\partial G(\gamma_f, \gamma_s)}{\partial \gamma'} = \begin{pmatrix} \frac{-exp(\gamma^f)}{1+r-exp(\gamma^f)} & 0 \\ 0 & t^* + \frac{-exp(\gamma^s)}{1+r-exp(\gamma^s)} \end{pmatrix}
$$

yields approximation of the covariance matrix

$$
Var\left(G(\gamma_f, \gamma_s)\right) \approx \begin{pmatrix} \left(\frac{-exp(\mu_\gamma)}{1+r-exp(\mu_\gamma)}\right)^2 \sigma_\gamma^2 & \frac{-exp(\mu_\gamma)\sigma_\gamma^2 \rho}{1+r-exp(\mu_\gamma)}\left(t^* + \frac{-exp(\mu_\gamma)}{1+r-exp(\mu_\gamma)}\right) \\ \frac{-exp(\mu_\gamma)\sigma_\gamma^2 \rho}{1+r-exp(\mu_\gamma)}\left(t^* + \frac{-exp(\mu_\gamma)}{1+r-exp(\mu_\gamma)}\right) & \left(t^* + \frac{-exp(\mu_\gamma)}{1+r-exp(\mu_\gamma)}\right)^2 \sigma_\gamma^2 \end{pmatrix}
$$

This covariance can be evaluated for specific choices of μ_γ, σ_γ^2, ρ and r. The underlying income formation model is quite artificial (infinite lifetimes), but simulation results reported in Stuhler (2010) verify that the size of the life-cycle bias under log-linear income formation is also non-zero for finite lifetimes.

---

[48]From $\lambda_{s,t} = \frac{Cov(y_i^*, y_{it})}{Var(y_i^*)} \cong \frac{\sigma_\eta^2 + t\sigma_\gamma^2/r}{\sigma_\eta^2 + \sigma_\gamma^2/r^2}$ it follows that t^* can be approximated by $1/r$.

A.3 Life-Cycle Bias: Right-Side Measurement Error

Assume that we wish to estimate the regression model (1), but that log lifetime income of fathers $y_{f,i}^*$ is approximated by $y_{f,it}$, log annual income at age t. Sons' log lifetime income $y_{s,i}^*$ is observed. We express the linear projection of $y_{f,it}$ on $y_{f,i}^*$ as

$$y_{f,it} = \lambda_{f,t} y_{f,i}^* + u_{f,it}$$

The probability limit of the OLS estimator of a linear regression of $y_{s,i}^*$ on $y_{f,it}$ is then

$$plim\,\hat{\beta}_t = \frac{Cov(y_{f,it}, y_{s,i}^*)}{Var(y_{f,it})}$$

$$= \theta_{f,t}\beta + \theta_{f,t}\frac{Cov(u_{f,it}, y_{s,i}^*)}{\lambda_{f,t} Var(y_{f,i}^*)}$$

where the first step follows since the errors ϵ_i and $u_{f,it}$ are by construction uncorrelated with $y_{f,i}^*$, and $\theta_t = \lambda_{f,t} Var(y_{f,i}^*)/\lambda_{f,t}^2\left(Var(y_{f,i}^*) + Var(u_{f,it})\right)$ is the "reliability ratio" that can be interpreted as the slope coefficient in the "reverse regression" of $y_{f,i}^*$ on $y_{f,it}$. The reliability ratio reduces to the familiar attenuation bias if $y_{f,it}$ is measured at age t^* such that $\lambda_{f,t} = 1$. The GEiV model is based on the assumption that $u_{f,it}$ is uncorrelated to $y_{s,i}^*$. It can account for the reliability ratio, but not for the remaining life-cycle bias in the second term of the last line that stems from correlation in the shape of income profiles within families.

A.4 Life-Cycle Bias: Left- and Right-Side Measurement Error

Assume that we wish to estimate the regression model (1), but that log lifetime incomes of fathers $y_{f,i}^*$ and sons $y_{f,i}^*$ are not observed and thus approximated by $y_{f,it}$ and $y_{s,it}$, log annual incomes at age t.[49] We express the linear projection of $y_{f,it}$ on $y_{f,i}^*$ as

$$y_{f,it} = \lambda_{f,t} y_{f,i}^* + u_{f,it}$$

and the linear projection of $y_{s,it}$ on $y_{s,i}^*$ as

$$y_{s,it} = \lambda_{s,t} y_{s,i}^* + u_{s,it}$$

The probability limit of the OLS estimator of a linear regression of $y_{s,it}$ on $y_{f,it}$ is then

$$plim\,\hat{\beta}_t = \frac{Cov(y_{s,it}, y_{f,it})}{Var(y_{f,it})}$$

$$= \frac{\beta\lambda_{s,t}\lambda_{f,t}Var(y_{f,i}^*) + \lambda_{f,t}Cov(u_{s,it}, y_{f,i}^*) + \lambda_{s,t}Cov(y_{s,i}^*, u_{f,it}) + Cov(u_{s,it}, u_{f,it})}{\lambda_{f,t}^2 Var(y_{f,i}^*) + Var(u_{f,it})}$$

If log annual income of fathers and sons are measured at ages such that $\lambda_{s,t} = \lambda_{f,t} = 1$ then the probability limit of the estimator reduces to

$$plim\,\hat{\beta}_t = \frac{\beta Var(y_{f,i}^*) + Cov(u_{s,it}, y_{f,i}^*) + Cov(y_{s,i}^*, u_{f,it}) + Cov(u_{s,it}, u_{f,it})}{Var(y_{f,i}^*) + Var(u_{f,it})}$$

an expression akin (except for the subscript t) to the general eq. (2).

---

[49]Note that for notational simplicity we here do not distinguish the age subscripts for fathers and sons.

References

ABOWD, J., AND D. CARD (1989): "On the Covariance Structure of Earnings and Hours Changes," *Econometrica: Journal of the Econometric Society*, 57(2), 411–445.

ATKINSON, A. (1980): "On Intergenerational Income Mobility in Britain," *Journal of Post Keynesian Economics*, 3(2), 194–218.

BAKER, M. (1997): "Growth-Rate Heterogeneity and the Covariance Structure of Life-Cycle Earnings," *Journal of Labor Economics*, 15(2), 338–75.

BECKER, G., AND N. TOMES (1979): "An Equilibrium Theory of the Distribution of Income and Intergenerational Mobility," *The Journal of Political Economy*, 87(6), 1153–1189.

———— (1986): "Human Capital and the Rise and Fall of Families," *Journal of Labor Economics*, 4(3), 1–39.

BHULLER, M., M. MOGSTAD, AND K. G. SALVANES (2011): "Life-Cycle Bias and the Returns to Schooling in Current and Lifetime Earnings," Unpublished working paper.

BJÖRKLUND, A. (1993): "A Comparison between Actual Distributions of Annual and Lifetime Income: Sweden 1951-89," *Review of Income and Wealth*, 39(4), 377–386.

BJÖRKLUND, A., AND M. JÄNTTI (1997): "Intergenerational Income Mobility in Sweden Compared to the United States," *American Economic Review*, 87(5), 1009–18.

BJÖRKLUND, A., AND M. JÄNTTI (2009): "Intergenerational Income Mobility and the Role of Family Background," in *Oxford Handbook of Economic Inequality*, ed. by B. N. Salverda Wiemer, and T. M. Smeeding. Oxford University Press.

BJÖRKLUND, A., M. JÄNTTI, AND M. J. LINDQUIST (2009): "Family Background and Income During the Rise of the Welfare State: Brother Correlations in Income for Swedish Men born 1932-1968," *Journal of Public Economics*, 93(5-6), 671 – 680.

BJÖRKLUND, A., M. LINDAHL, AND E. PLUG (2006): "The Origins of Intergenerational Associations: Lessons from Swedish Adoption Data," *Quarterly Journal of Economics*, 121(3), 999–1029.

BLACK, S. E., AND P. DEVEREUX (2011): "Recent Developments in Intergenerational Mobility," in *Handbook of Labor Economics*, ed. by O. Ashenfelter, and D. Card, vol. 4A/4B. Elsevier.

BLUNDELL, R., AND I. PRESTON (1998): "Consumption Inequality and Income Uncertainty," *Quarterly Journal of Economics*, 113(2), 603–640.

BÖHLMARK, A., AND M. LINDQUIST (2006): "Life-Cycle Variations in the Association between Current and Lifetime Income: Replication and Extension for Sweden," *Journal of Labor Economics*, 24(4), 879–896.

BROWNING, M., M. EJRNÆS, AND J. ALVAREZ (2010): "Modelling Income Processes with Lots of Heterogeneity," *Review of Economic Studies*, 77(4), 1353–1381.

CORAK, M., AND P. PIRAINO (2010): "The Intergenerational Transmission of Employers," *Journal of Labor Economics*, 29(1), 37–68.

COUCH, K., AND D. LILLARD (1998): "Sample Selection Rules and the Intergenerational Correlation of Earnings," *Labour Economics*, 5(3), 313–329.

CUNHA, F., J. HECKMAN, AND S. NAVARRO (2005): "Separating Uncertainty from Heterogeneity in Life Cycle Earnings," *Oxford Economic Papers*, 57(2), 191.

DAHL, M., AND T. DELEIRE (2008): "The Association Between Children's Earnings And Fathers' Lifetime Earnings: Estimates Using Administrative Data," Discussion Paper No. 1342-08, Institute for Research on Poverty, University of Wisconsin-Madison.

FRIEDMAN, M. (1957): *A Theory of the Consumption Function*. National Bureau of Economic Research.

GOUSKOVA, E., N. CHITEJI, AND F. STAFFORD (2010): "Estimating the Intergenerational Persistence of Lifetime Earnings with Life Course Matching: Evidence from the PSID," *Labour Economics*, 17(3), 592–597.

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

GUVENEN, F. (2007): "Learning Your Earning: Are Labor Income Shocks Really Very Persistent?," *American Economic Review*, 97(3), 687–712.

——— (2009): "An Empirical Investigation of Labor Income Processes," *Review of Economic Dynamics*, 12(1), 58–79.

HAIDER, S., AND G. SOLON (2006): "Life-Cycle Variation in the Association between Current and Lifetime Earnings," *American Economic Review*, 96(4), 1308–1320.

HECKMAN, J. J., L. J. LOCHNER, AND P. E. TODD (2006): "Earnings Functions, Rates of Return and Treatment Effects: The Mincer Equation and Beyond," in *Handbook of the Economics of Education*, ed. by E. Hanushek, and F. Welch, vol. 1, chap. 7, pp. 307–458. Elsevier.

JENKINS, S. (1987): "Snapshots versus Movies: 'Lifecycle biases' and the Estimation of Intergenerational Earnings Inheritance," *European Economic Review*, 31(5), 1149–1158.

——— (2009): "Spaghetti Unravelled: A Model-Based Description of Differences in Income-Age Trajectories," ISER Working Paper Series 2009-30, Institute for Social and Economic Research, University of Essex.

LILLARD, L. A., AND Y. WEISS (1979): "Components of Variation in Panel Earnings Data: American Scientists, 1960-70," *Econometrica*, 47(2), 437–54.

MACURDY, T. (1982): "The Use of Time Series Processes to Model the Error Structure of Earnings in a Longitudinal Data Analysis," *Journal of Econometrics*, 18(1), 83–114.

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

MEGHIR, C., AND L. PISTAFERRI (2004): "Income Variance Dynamics and Heterogeneity," *Econometrica*, 72(1), 1–32.

MOCETTI, S. (2007): "Intergenerational Earnings Mobility in Italy," *The B.E. Journal of Economic Analysis & Policy*, 7(2), Article 5.

MOFFITT, R., AND P. GOTTSCHALK (1995): "Trends in the Covariance Structure of Earnings in the United States: 1969-1987," Institute for Research on Poverty Discussion Papers 1001-93, University of Wisconsin Institute for Research on Poverty.

NICOLETTI, C., AND J. ERMISCH (2007): "Intergenerational Earnings Mobility: Changes across Cohorts in Britain," *The B.E. Journal of Economic Analysis & Policy*, 7(2), Article 9.

NILSEN, Ø., K. VAAGE, A. AAKVIK, AND K. JACOBSEN (forthcoming): "Intergenerational Earnings Mobility Revisited: Estimates Based on Lifetime Earnings," *Scandinavian Journal of Economics*.

PIRAINO, P. (2007): "Comparable Estimates of Intergenerational Income Mobility in Italy," *The B.E. Journal of Economic Analysis & Policy*, 7(2), Article 1.

RAAUM, O., B. BRATSBERG, K. RØED, E. ÖSTERBACKA, T. ERIKSSON, M. JÄNTTI, AND R. NAYLOR (2007): "Marital Sorting, Household Labor Supply, and Intergenerational Earnings Mobility across Countries," *The B.E. Journal of Economic Analysis & Policy*, 7(2), Article 7.

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

SOLON, G. (1999): "Intergenerational Mobility in the Labor Market," in *Handbook of Labor Economics*, ed. by O. Ashenfelter, and D. Card, vol. 3A, chap. 29, pp. 1761–1800. Elsevier.

STUHLER, J. (2010): "Empirical Strategies to Eliminate Life-Cycle Bias in the Intergenerational Elasticity of Earnings Literature," IZA Discussion Papers 5245, Institute for the Study of Labor (IZA).

VOGEL, T. (2006): "Reassessing Intergenerational Mobility in Germany and the United States: The Impact of Differences in Lifecycle Earnings Patterns," SFB 649 Discussion Papers 2006-055, Sonderforschungsbereich 649, Humboldt University, Berlin, Germany.

ZIMMERMAN, D. J. (1992): "Regression toward Mediocrity in Economic Stature," *American Economic Review*, 82(3), 409–29.