Institute for Research on Poverty
Discussion Paper No. 1342-08

The Association between Children's Earnings and Fathers' Lifetime Earnings: Estimates Using Administrative Data

Molly Dahl
Congressional Budget Office
U.S. Congress
E-mail: molly.dahl@cbo.gov

Thomas DeLeire
La Follette School of Public Affairs
University of Wisconsin–Madison
E-mail: deleire@wisc.edu

August 2008

We thank John Abowd, Dean Lillard, Len Lopoo, Erzo Luttmer, Bhash Mazumder, Gary Solon and seminar participants at the Congressional Budget Office, the Brookings Institution, the Urban Institute, the Institute for Research on Poverty Summer Workshop, the University of Kentucky, Virginia Commonwealth University, and the Society for Labor Economists Annual Meeting for helpful comments. All errors are our own. The views presented in this paper are those of the authors and do not reflect those of the Congressional Budget Office.

Abstract

Knowledge of the degree of intergenerational mobility in an economy is essential for assessing the fairness of the earnings distribution. In this paper, we provide estimates of the degree of intergenerational mobility in the United States using administrative earnings data from the Social Security Administration's records. These data contain nearly career-long earnings histories for a large sample of U.S. fathers, and their children's earnings around an age that is likely to be a good proxy for lifetime earnings. We examine two different measures of mobility: (1) the association between fathers' and children's log earnings (the intergenerational elasticity or IGE) and (2) the association between fathers' and children's relative positions in their respective earnings distributions (or the intergenerational rank association or IRA). We show that estimates of the IGE are quite sensitive to choice of specification and sample and range from 0.26 to 0.63 for sons and from 0 to 0.27 for daughters. That is, a 10 percent increase in fathers' earnings is associated with a 3 percent to 6 percent increase in sons' earnings and a 0 percent to 3 percent increase in daughters' earnings. By contrast, our estimates of the IRA are robust to both specification and sample choices and show that a 10 percentile increase in a father's relative position is associated with roughly a 3 percentile increase in his son's and roughly a 1 percentile increase in his daughter's relative earnings positions. Nonparametric estimates of the IRA show relatively more immobility among the children of men below the 10^{th} percentile and above the 80^{th} percentile of lifetime earnings.

Key Words: Intergenerational, Mobility, Earnings

The Association between Children's Earnings and Fathers' Lifetime Earnings: Estimates Using Administrative Data

I. INTRODUCTION

Knowledge of the degree of intergenerational mobility in an economy is essential for assessing the fairness of the earnings distribution. If there are opportunities for individuals of all backgrounds to achieve high levels of lifetime earnings (that is, if the intergenerational mobility in earnings is high), even an economy with a wide earnings distribution might be considered fair. In fact, the degree of intergenerational mobility in earnings has been often associated with the amount of economic opportunity in an economy.

A large and growing literature has examined a particular measure of intergenerational mobility in the United States, the intergenerational elasticity (IGE) between fathers' and children's earnings, and has produced a wide range of estimates. The earliest literature (reviewed by Becker and Tomes, 1986) found evidence of the United States being a highly mobile society (with estimates of the IGE in earnings of roughly 0.2). However, estimates of the IGE based on national longitudinal samples range from roughly 0.1 to 0.5, with a typical estimate being 0.4 (Solon, 1999) suggesting that the actual degree of intergenerational mobility is lower. A more recent paper, Mazumder (2005), used long panels of earnings from the Social Security Administration (SSA) and found an even greater association between the earnings of fathers and their children (with an IGE of 0.6), and suggested that the United States is substantially less mobile than previously believed.

A related literature has examined the intergenerational association of occupational ranking. For example, Blau and Duncan (1967) find an association of about 0.45 while, according to a recent review by Beller and Hout (2006), current estimates are about 0.30.

In this paper, we use high-quality administrative earnings data on a large cohort of fathers and their adult sons and daughters.[1] First, we provide estimates of IGE in earnings between fathers' lifetime earnings and the earnings of their sons and daughters. We measure children's average earnings around an age—36—that likely acts as an unbiased proxy for their lifetime earnings (Haider and Solon, 2006). We also use nearly career-long earnings histories of fathers—averaged over the ages of 20 to 55—which should eliminate both attenuation bias from measurement error and life-cycle bias from the measurement of fathers' average earnings over only certain age ranges. This specification of children's and fathers' earnings represents a departure from much of the previous literature, which typically averages fathers' and children's earnings over the available years of data and adjust for age using flexible controls.[2] Importantly, we then document how sensitive the base specification is to specification and sample choice. In particular, we show that the estimate of the intergenerational elasticity between fathers' and sons' lifetime earnings ranges from 0.26 to 0.63 depending upon specification and sample. Estimates from samples that include fathers who have years with no labor market earnings tend to be lower, indicating higher mobility. Estimates from specifications that control for life-cycle and attenuation biases tend to be larger, indicating lower levels of mobility.

Second, we provide estimates of an alternative measure of intergenerational mobility—what we call the intergenerational rank association (IRA)—which is the association between fathers' positions in the lifetime earnings distribution and their children's positions in the earnings distribution. Unlike the IGE, these estimates have the feature of being highly robust across specifications and samples, including those with fathers that have years of zero earnings. Our estimates show that a 10 percentile increase in a

---

[1]Only a few previous papers have estimated the association between fathers' and daughters' earnings, for example, Chadwick and Solon (2002).

[2]Grawe (forthcoming) is an exception. However, because he uses single-year measures of earnings when estimating the IGE in earnings to more clearly emphasize the existence of life-cycle bias, his estimates of the IGE in earnings likely are attenuated (and are substantially lower than those found in the recent literature). Vogel (2007) also explores how life-cycle bias may affect estimates of intergenerational mobility in Germany and in the U.S.

father's position is associated with his son being roughly 3 percentiles higher and with his daughter being roughly 1 percentile higher in their respective earnings distributions.

Third, nonparametric estimates of the IRA reveal relatively more mobility in the broad middle of fathers' lifetime earnings distribution (between the 10th and 80th percentiles) and relatively more immobility at both the bottom and the top.

This paper proceeds as follows. Section 2 reviews the literature and discusses some of the methods used and measurement issues identified by previous studies. Section 3 describes the data. Section 4 describes our methods. Section 5 presents estimates of the IGE and IRA across a range of specifications and samples. Section 6 presents several nonparametric estimates of intergenerational mobility. Section 7 summarizes our conclusions.

II. METHOD AND MEASUREMENT ISSUES

In this section, we review the literature on intergenerational mobility focusing on method and measurement issues. First, we review the extensive literature estimating the intergenerational elasticity. Second, we review the literature providing estimates related to our second measure, the intergenerational rank association, including papers which estimate transition matrices and papers estimating the degree of association between the occupations of fathers and sons.

The Intergenerational Elasticity

There is a large literature in both economics and sociology providing estimates of the intergenerational elasticity in earnings. The early studies of the IGE in earnings (surveyed by Becker and Tomes, 1986) tend to regress single-year measures of sons' earnings on single-year measures of fathers' earnings (along with flexible controls for age), as in the following model:

$$y_{it} = \alpha + \beta x_{it} + f\left(\gamma, Age_i\right) + \varepsilon_{it} . \quad (1)$$

These studies found relatively small estimates of the IGE in earnings (β), which range around 0.2. However, due to transitory fluctuations in earnings and measurement error, these estimates likely are subject to attenuation bias (Solon 1992, 1999).[3]

To partially address this issue, a large number of studies subsequently used nationally representative data from the Panel Study of Income Dynamics (PSID) and the National Longitudinal Survey (NLS).[4] These studies typically averaged fathers' earnings over a 3- to 5-year period so as to reduce measurement error and attenuation bias. Of the 15 studies surveyed in Solon (1999), 3 estimate the IGE in earnings to be around 0.2 and 12 studies estimate it to be between 0.3 and 0.5. The studies with low estimates of the IGE tend to use samples of relatively young men, which is likely related to life-cycle bias as we discuss below. These studies also differed in how they treated observations with low, missing, or zero annual earnings in some years. Some studies (e.g., Zimmerman, 1992; Solon, 1992) exclude sons and fathers who usually work less than 30 hours per week or who work less than 30 weeks per year and also exclude observations with zero or missing earnings in any year. Others (e.g., Couch and Dunn, 1997; Eide and Showalter, 1999; Couch and Lillard, 1998) exclude years with missing data but include years with zero earnings when computing the averages and do not exclude part-time workers. Studies that include years with zero earnings and part-time workers tend to find lower estimates of the IGE in earnings than studies that do not.

As noted in previous studies, even estimating Equation 1 using 5-year averages of fathers' earnings will likely lead to some attenuation bias, especially in the presence of serially correlated transitory shocks to earnings. To address this issue, Mazumder (2005) uses Social Security Administration earnings data to calculate average father's earnings over many years of data. His estimates using just 2 years of fathers' earnings data are about 0.25; those using 4 to 7 years of data are about 0.3 to

---

[3]Solon (1999) notes that because the data sets used in these early analyses tended to have fairly homogenous samples, the attenuation bias is likely to be even greater.

[4]A notable study which does not use survey data is Corak and Heisz (1999), who use administrative data from Canada and estimate the IGE in earnings to be roughly 0.2.

0.5 (roughly equal to the estimates in most studies); while those using 10 or 16 years of data are much larger, 0.55 to 0.6. These results confirm that attenuation bias is an issue and suggest that "the United States is substantially less mobile than previous research indicated" (Mazumder, 2005, p. 235).

In an important paper, Haider and Solon (2006) note that attenuation bias is only one problem in estimating the IGE in earnings. Single-year measures of earnings do not follow classical errors in variables. Instead, the relationship between single-year measures of earnings and lifetime earnings depends upon the age that the single year of earnings is measured. Because individuals with higher lifetime earnings also tend to experience rapid earnings growth when young, differences tend to be underestimated at younger ages and overestimated at older ages.

Haider and Solon (2006) and Bohlmark and Lindquist (2006) use administrative data to calculate lifetime earnings from samples of men from the United States born between 1931 and 1933 and men and women from Sweden born between 1929 and 1933. They find that current earnings best proxy lifetime earnings between roughly the ages of 32 and 40. Because most researchers have tended to assume classical errors in variables, they have not worried about using only single-year measures of children's earnings or the age at which children's earnings are measured. This life-cycle bias may lead to underestimates of the IGE in earnings when using children's earnings measured in their 20s, as many of the studies surveyed by Solon (1999) had done. This insight led Haider and Solon to speculate that "many estimates of the intergenerational earnings elasticity have been subject to substantial attenuation inconsistency from left-side measurement error in addition to the well-known inconsistency from right-side measurement error" (p. 1319) (though neither Haider and Solon nor Bohlmark and Lindquist provide new estimates of the IGE that correct for life-cycle bias). In fact, life-cycle bias may explain some of the variation across studies in estimates of the IGE in Solon's (1999) survey (though, as noted above, the samples used in these studies also differed in other dimensions such as the inclusion of observations with years with zero earnings). In the analysis in this paper, we measure sons' and daughters' earnings around age 36 to avoid bias from life-cycle measurement error.

Life-cycle bias does not just affect the measurement of children's lifetime earnings. It also affects the measurement of fathers' lifetime earnings, even when multi-year averages of earnings are used. In fact, the use of multi-year averages will reduce noise but is still subject to life-cycle bias and could lead to either amplification bias or attenuation bias (Haider and Solon 2006). A solution to this problem is to use lifetime earnings of fathers when possible. In the analysis presented in this paper, we use nearly career-long averages of fathers' earnings. This solution deals with both "right-side" problems associated with the mismeasurement of fathers' earnings: attenuation bias and life-cycle bias.

In summary, there are three methodological issues highlighted in the literature on the intergenerational transmission of earnings: (1) attenuation bias from right-side measurement error from using single-year or multi-year averages of fathers' earnings as a proxy for lifetime earnings; (2) right-side life-cycle bias (which can result in either attenuation or amplification bias) when using either a single-year or a multi-year average of fathers' earnings as a proxy for lifetime earnings; and (3) attenuation or amplification bias from left-side life-cycle bias from using single-year or multi-year averages of sons' or daughters' earnings as a proxy for lifetime earnings. We address the first two issues by using a lifetime measure of fathers' earnings. We address the third issue by measuring sons' and daughters' earnings as averages centered on age 36.

The Intergenerational Association in Ranks

A large number of studies have documented the extent of intergenerational mobility by calculating transition matrices (e.g., Peters, 1992; Zimmerman, 1992; Isaacs, 2008). The pictures of mobility across generations created by these transition matrices are more complex than the single number estimates of the IGE in earnings. A transition matrix shows the probability of a child's being in a given quintile (quartile, or decile, depending upon the study) of the earnings distribution conditional upon his or her father's position in the earnings distribution.

The estimates of transition matrices from different studies are often remarkably similar, even when those studies yield very different estimates of the IGE. For example, Peters (1992) and Zimmerman

(1992) both use the NLS to estimate transition matrices. Interestingly, while these two studies present widely different estimates of the IGE in earnings for fathers and sons—0.14 in Peters (1992) and 0.54 in Zimmerman (1992)—most likely because they make different choices regarding their samples and specifications (e.g., inclusion of part-time workers, workers with zero earnings in some years, and age of the sons)—their transition matrices are remarkably similar. This similarity in transition matrices across studies suggests that the methodological and sample choice issues that may have led to the wide range of estimates in the literature of the IGE in earnings may be less of an issue in estimates of transition matrices.

A related literature has examined the intergenerational association in the occupations of fathers and sons. These studies, (e.g., Blau and Duncan, 1967) often construct an ordered ranking of occupational status prestige and then regress the rank of the son on the rank of the father. In the United States, these studies have typically found that the association between fathers' and sons' occupational rankings is roughly 0.3 to 0.4 (Beller and Hout, 2006). While the degree of mobility will likely depend in part on the method used to rank occupations, there is a greater consensus across these studies than across studies using the IGE.

In this paper, we adopt a method that is very similar to that used in the literature on transition matrices and the intergenerational association in occupational rankings—we measure the association between fathers' and children's rank in the earnings distribution. We are unaware of any published studies that have adopted this measure of mobility across generations.

Measurement Issues

There are a number of measurement issues that have confronted researchers seeking to measure the extent of intergenerational earnings mobility. These measurement issues include (1) how to treat years of zero earnings, (2) whether to include part-time workers, (3) how to deal with top-coded earnings, and (4) how to deal with earnings that are not covered by Social Security (when using administrative data). Not surprisingly, studies differ on how they handle each of these issues.

First, studies that estimate the IGE in earnings using single-year measures of log fathers' and log children's earnings effectively drop observations with zero annual earnings. Solon (1992) and Zimmerman (1992) average over several years of fathers' earnings and drop observations with zero earnings in any of those years. Other studies make different choices. For example, Couch and Dunn (1997) and Eide and Showalter (1999) include years of zero earnings for both sons and fathers when constructing log average earnings (dropping observations with zero earnings in every year). Peters (1992), alternatively, excludes years of zero earnings when constructing multi-year averages of earnings (that is, only averages over the positive years of earnings). Perhaps as a result, Peters (1992), Eide and Showalter (1999), and Couch and Dunn (1997) estimate relatively low IGEs in earnings, or relatively high levels of mobility, (ranging from 0.15 to 0.34), while Solon (1992) and Zimmerman (1992) estimate relatively high IGEs, or relatively low levels of mobility (0.4 to 0.5). When Couch and Dunn (1997) exclude observations with years of zero earnings, their estimate increases substantially. Couch and Lillard (1998) argue that estimates from the PSID and NLS are sensitive to the decision whether to exclude observations with years of zero earnings (and Solon [1998] documents that this sensitivity is, in fact, itself sensitive and is driven by a few observations). In this paper, we show that our estimates of the IGE in earnings are sensitive to this choice (though our estimates of the association between fathers' and children's relative earnings positions are not).

Second, some papers (Solon, 1992; Zimmerman, 1992) restrict their samples to full-time, full-year workers. The use of administrative data on annual earnings precludes such a restriction because information on hours and weeks worked is not available.

Third, the papers cited above that use SSA administrative data (Haider and Solon, 2006; and Mazumder, 2005) have to deal with issues related to the fact that their data are top-coded at the Social Security taxable maximum. As a result, top-coded earnings must be imputed.[5] In this paper, we use, when

[5]For example, Mazumder (2005) imputes top-coded earnings for fathers using CPS averages by race and education. For children, he also imputes top-coded earnings and non-covered earnings.

possible, uncensored and complete earnings data from SSA, which are available from 1984 onwards. Earnings data prior to 1984 that we use to construct lifetime earnings are potentially top-coded. For these years of earnings data, we impute top-coded earnings using an individual's earnings history. Details of this imputation procedure are described in the Appendix.

Fourth, some SSA administrative datasets only include earnings covered by Social Security. For example, Haider and Solon (2006) use earnings that are covered by Social Security in their main analyses while Mazumder (2005) uses only covered earnings in all of his analyses. In this paper, we again use uncensored and complete earnings data from SSA, which are available from 1984 onwards and which include non-covered earnings. Prior to 1984, we must use covered earnings as previous researchers do. We are able, however, to examine whether our results are sensitive to dropping fathers who likely had substantial amounts of non-covered earnings prior to 1984 (fathers who ever were in the armed forces and, alternatively, by dropping fathers who reported being either self-employed or working for the federal, state, or local government in 1984).

III. DATA

Our data come from the 1984 Survey of Income and Program Participation (SIPP) matched to Social Security Administration's (SSA) detailed earnings records (DER) and summary earnings records (SER). The 1984 SIPP is a nationally representative sample of households that were initially interviewed between October 1983 and January 1984. The SER contain annual earnings for workers that are top-coded at the Social Security taxable maximum since 1951. The DER data, by contrast, are not top-coded but are only available since 1984. We restrict our sample to men who were born between 1931 and 1949 (so that we have complete earnings histories from age 20 until age 55) and who lived with at least one of their own children who were between the ages of 12 and 21 as of the first interview (born between 1963 and 1972).[6] The sample is necessarily further restricted to those fathers and children who provided Social Security numbers to the Census interviewers and who were successfully matched to the SSA data. Ninety-five percent of fathers were successfully matched, but only 75 percent of children were matched. Our final sample consists of 1,869 father-son pairs and 1,652 father-daughter pairs. In specification checks that determine whether our results are sensitive to non-covered earnings, reported below, we also restrict our sample to fathers who never served in the armed forces and, alternatively, to those who were not self-employed and did not work for federal, state, or local governments in 1984. We use all siblings who meet the sample restrictions and adjust our standard errors accordingly.[7]

The use of SSA earnings records enables us to construct a lifetime earnings measure for a large sample of fathers.[8] Our measure of lifetime earnings is the natural logarithm of average annual earnings from age 20 and age 55, including years of zero earnings (all earnings data are inflated to 2005 dollars using the CPI-U-RS). Below, we determine how robust our results are to the choice of including years of zero earnings in our measure of lifetime earnings. Alternative choices we examine include (a) restricting our sample of fathers to those with positive earnings in every year from age 25 to 55, (b) restricting our sample of fathers to those with positive earnings in at least 16 years, (c) measuring lifetime earnings as the natural logarithm of average positive years of annual earnings from age 20 and age 55, and (d) measuring lifetime earnings as the natural logarithm of average positive years of annual earnings from the first age at which continuous employment began (which we define as the beginning of at least five consecutive years of positive earnings) to age 55. For all of the fathers in our sample, we observe annual earnings in each year from age 20 to 55, spanning the years 1951 to 2004.

---

[6]As we demonstrate below, the results are not sensitive to imposing more narrow restrictions on the ages of the children in our sample.

[7]All results in the paper are unweighted. Conducting the analyses using the 1984 SIPP sample weights does not affect the results.

[8]Social Security earnings data have the additional feature of likely being more accurate than earnings reported in surveys (Bound and Krueger 1991).

Prior to 1984, the only administrative earnings data available come from the SER, which are limited in two ways. First, they are top-coded at the maximum amount subject to the Social Security tax; this amount varies by year. Second, the SER earnings are available only for jobs covered by Social Security. Over 70 percent of earnings were covered between 1951 and 1956, and over 80 percent between 1957 and 1984. In the specifications in which we restrict our sample to fathers with positive earnings at every age, we are also effectively restricting our sample to fathers in jobs covered by Social Security.[9] After 1984, we use data from the DER, which are not top-coded and represent earnings from all jobs (including self-employment).[10]

The availability of the DER yields a large sample of observations with uncensored earnings data. These data enable us to impute top-coded data using each observation's very rich earnings history. In particular, we use propensity score methods to match top-coded observations in the SER to uncensored observations from the DER. Because the censoring point varies by year, we implement this matching procedure separately for each age and year for which we have top-coded observations. The matching variables we use include workers' annual positions in the earnings distribution for the 5 years prior to and following the top-coded observation, as well as a set of corresponding indicator variables for whether his or her earnings in those years would also have been top-coded. For top-coded observations at ages under age 25 (over age 50), we use the 10 observations on annual earnings between ages 20 and 29 (46 and 55) as matching variables. We then impute the percentile in the earnings distribution for each top-coded observation using a nearest neighbor matching procedure. Details are provided in the Appendix.

---

[9]Mazumder (2005) also does this.

[10]The DER data are available beginning in 1978. Because of concerns regarding the reliability of the data in the first few years of collection, we use the SER to measure earnings prior to 1984 and the DER to measure earnings thereafter.

IV. METHODS

In this section we outline the methods we use in estimating the IGE and the IRA in earnings between our sample fathers and their children.

Intergenerational Elasticity in Earnings

We estimate the following model to determine the IGE in earnings:

$$y_{it} = \alpha + \beta x_{it} + \varepsilon_{it} \tag{2}$$

where:

y_{it} is child's earnings; and

x_{it} is father's lifetime earnings.

We measure both child's earnings and father's earnings in several different ways. In general, we use a lifetime measure of father's earnings and measure children's earnings at the same ages in a given specification; we do not need to make age adjustments as in Equation 1. We use the log of average of earnings from age 35 to 37 as our proxy for children's lifetime earnings and the log of average earnings from age 20 to 55 for father's lifetime earnings in our baseline specification.[11]

Haider and Solon (2006) and Bohlmark and Lindquist (2006) determine the extent to which annual earnings can proxy for lifetime earnings for cohorts of American and Swedish men. Haider and Solon (2006) find that, for their cohort of men born between 1931 and 1933, annual earnings between ages 32 and 40 are good proxies for lifetime earnings. Bohlmark and Lindquist find that, for their cohort of men born between 1929 and 1933, annual earnings between ages 34 and 40 are good proxies for lifetime earnings. While Haider and Solon do not determine the extent of life-cycle biases for women,

[11] 85 sons (8 percent) and 145 daughters (15 percent) are dropped from the sample because they do not have earnings between the ages of 35 and 37.

Bohlmark and Lindquist find that ages 30 to 33 and age 35 are good proxies for the lifetime earnings of women.

Because the cohorts used in each of Haider and Solon (2006) and Bohlmark and Lindquist (2006) studies are slightly older than the fathers in our sample (and substantially older than our sample of sons and daughters), we replicate their findings using three samples: (1) our sample of fathers; (2) a cohort of men born between 1945 and 1949; and (3) a cohort of women born between 1945 and 1949 (the youngest cohort for which we could construct lifetime earnings).[12] In particular, we regress annual earnings at each age between 20 and 55 on a measure of lifetime earnings. The coefficient on lifetime earnings determines the extent of life-cycle biases. If that coefficient is close to one, then these biases will be small, which it is at about age 38 (see Appendix Figure A1). We also regress multi-year averages of earnings between various age ranges (see Appendix Table A1). From these procedures, it appears that earnings averaged from age 35 to 37 are a reasonable proxy for lifetime earnings for our sample of fathers and for both men and women.[13]

Intergenerational Rank Association

Our second measure of intergenerational mobility is the IRA—the association between children's and fathers' relative positions in their respective earnings distributions. To create this measure, we regress children's earnings percentile (measured at ages 35 through 37) in their own gender's earnings distribution on fathers' lifetime earnings percentile. This measure is similar to those relating the

[12] We impute top-coded earnings for years prior to 1984 and use the non-censored data available from 1984 onwards and follow Bjorklund (1993) and Bohlmark and Lindquist (2006). Haider and Solon (2006) employ a complex limited dependent variable model to deal with censoring. Our results do not differ appreciably from those of Haider and Solon.

[13] We also have used our data to replicate Haider and Solon's (2006) "reverse regression" to measure the extent of attenuation bias that would result from using single-age measures of earnings. The results are available upon request and confirm that estimates using single-age measures of earnings, especially those measured at younger and older ages, would be severely attenuated. Attenuation bias is also present when using multiyear average of earnings.

occupational ranking of children to that of fathers (e.g., Beller and Hout, 2006). In particular, we estimate the following OLS regression,

$$p_{it}^{y} = \alpha + \beta p_{it}^{X} + \varepsilon_{it} \tag{3}$$

where:

p^{y}_{it} is the child's percentile of earnings in their own gender's earnings distribution (measured over an interval centered around age 36); and

p^{X}_{it} is the father's percentile of lifetime earnings (measured between age 20 and 55).

These measures impose a linear relationship between the position of children in the earnings distribution and the position of fathers in the lifetime earnings distribution and thus, like the IGE, yield a single number for the degree of intergenerational mobility. In order to allow for a potentially nonlinear and highly flexible relationship between children's and fathers' positions in their respective earnings distributions, we also calculate a nonparametric estimate of this relationship.

To do this, we create moving blocks (of five percentiles) of fathers ordered by their percentile in the lifetime earnings distribution. For example, the first block includes all fathers in the first through fifth percentiles of lifetime earnings; the second block includes fathers in the second through sixth percentiles; and the 96th block includes fathers in the 96th through 100th percentiles. For each block, we calculate the percentile in the overall children's earnings distribution at which the 20th percentile, median, and 80th percentile of earnings of the children in that block fall (where earnings of children once again are measured as the average between ages 35 and 37). We do this calculation separately for sons and daughters.

V. RESULTS

In this section, we present our estimates of the IGE and IRA in earnings and show how these estimates are sensitive to choice of specification and sample. Our evidence on the extent of intergenerational mobility based on estimates of the IGE and of the IRA can be summarized as follows:

- Estimates of the IGE in earnings range from 0.26 to 0.63 for sons and from 0 to 0.27 for daughters, being sensitive to small changes in sample and in the definition of fathers' lifetime earnings.
- Estimates of the IRA in earnings, by contrast, are robust across samples and definitions of fathers' lifetime earnings and are roughly 0.3 for sons and 0.1 for daughters.
- Estimates of both IGE and IRA are sensitive to the age at which children's earnings are measured; estimates based on the earnings of children measured in their 20s are very low and are likely biased down due to life-cycle bias. Estimates of the IGE are sensitive to the age spans over which fathers' earnings are measured while estimates of the IRA are not.
- Estimates of the IGE are sensitive to the number of years used to calculate fathers' lifetime earnings. Surprisingly, estimates of the IRA are not sensitive to this choice; this finding may be of use to researchers lacking access to the extremely long panels of earnings that we use to construct fathers' lifetime earnings.

Sensitivity of Estimates of the IGE and IRA in Earnings to Sample and to Definition of Fathers' Lifetime Earnings

Our first set of estimates of the IGE and IRA in earnings between fathers and sons is presented in Table 1a and that between fathers and daughters is presented in Table 1b. The estimates of the IGE in earnings presented in these tables are all based on our baseline specification (Equation 2) but differ from one another in how we define the sample or lifetime earnings for fathers. Similarly, the estimates of the IRA in earnings are based on Equation 3 and differ from one another based on the same set of changes in sample and lifetime earnings definitions. The dependent variable in all IGE models is the natural logarithm of average earnings from age 35 to 37 for children and is their position in the earnings distribution averaged from age 35 to 37 in all IRA models.

The estimates of the IGE for sons and fathers, presented in Table 1a, are all from the same specification but with slight differences in sample and in the definition of lifetime earnings for fathers (as we describe below) and range widely—from 0.26 to 0.63.

Table 1a
Estimates of the Intergenerational Elasticity and of the Intergenerational Rank Association between Sons' Earnings and Fathers' Lifetime Earnings

Model	(1) Intergenerational Elasticity (IGE) Coefficient (Standard error) *Number of Obs.*	(2) Intergenerational Rank Association (IRA) Coefficient (Standard error) *Number of Obs.*	(3) Sample	(4) Definition of Father's Lifetime Earnings
(1)	0.299 (0.069) *1,000*	0.292 (0.033) *1,000*	All fathers	Log of average earnings from age 20 to 55 including years of zero earnings
(2)	0.259 (0.080) *526*	0.314 (0.043) *526*	Fathers Never in the Armed Forces as of 1984	Log of average earnings from age 20 to 55 including years of zero earnings
(3)	0.280 (0.081) *700*	0.320 (0.039) *700*	Fathers Not Working in Government or Self-Employed Sectors in 1984	Log of average earnings from age 20 to 55 including years of zero earnings
(4)	0.498 (0.068) *994*	0.313 (0.034) *994*	All fathers	Log of average earnings beginning with the first 5 consecutive years of positive earnings to age 55
(5)	0.507 (0.082) *1,000*	0.323 (0.033) *1,000*	All fathers	Log of average of years of positive earnings from age 20 to 55
(6)	0.482 (0.074) *949*	0.293 (0.036) *949*	Fathers with 16 or more years of earnings	Log of average earnings from age 20 to 55 including years of zero earnings
(7)	0.632 (0.106) *516*	0.395 (0.055) *516*	Fathers with earnings at every age from 25 to 55	Log of average earnings from age 20 to 55 including years of zero earnings

Source: Authors' calculations from 1984 SIPP-SSA matched file.
Notes: The dependent variable is sons' log average earnings from age 35 to 37 (IGE) or sons' position in distribution of earnings averaged from age 35 to 37. Fathers' lifetime earnings are defined as indicated.
Standard errors are adjusted for the fact that some fathers appear in the sample more than once.

Table 1b
Estimates of the Intergenerational Elasticity and of the Intergenerational Rank Association between Daughters' Earnings and Fathers' Lifetime Earnings

Model	(1) Intergenerational Elasticity (IGE) Coefficient (Standard error) *Number of Obs.*	(2) Intergenerational Rank Association (IRA) Coefficient (Standard error) *Number of Obs.*	(3) Sample	(4) Definition of Father's Lifetime Earnings
(1)	0.177 (0.073) *820*	0.120 (0.037) *820*	All fathers	Log of average earnings from age 20 to 55 including years of zero earnings
(2)	0.117 (0.093) *422*	0.103 (0.049) *422*	Fathers Never in the Armed Forces as of 1984	Log of average earnings from age 20 to 55 including years of zero earnings
(3)	0.258 (0.082) *600*	0.165 (0.042) *600*	Fathers Not Working in Government or Self-Employed Sectors in 1984	Log of average earnings from age 20 to 55 including years of zero earnings
(4)	0.193 (0.072) *817*	0.158 (0.037) *817*	All fathers	Log of average earnings beginning with the first 5 consecutive years of positive earnings to age 55
(5)	0.269 (0.093) *820*	0.166 (0.037) *820*	All fathers	Log of average of years of positive earnings from age 20 to 55
(6)	0.215 (0.087) *777*	0.146 (0.039) *777*	Fathers with 16 or more years of earnings	Log of average earnings from age 20 to 55 including years of zero earnings
(7)	-0.041 (0.161) *427*	0.082 (0.063) *427*	Fathers with earnings at every age from 25 to 55	Log of average earnings from age 20 to 55 including years of zero earnings

Source: Authors' calculations from 1984 SIPP-SSA matched file.
Notes: The dependent variable is daughters' log average earnings from age 35 to 37 (IGE) or daughters' position in distribution of earnings averaged from age 35 to 37. Fathers' lifetime earnings are defined as indicated.
Standard errors are adjusted for the fact that some fathers appear in the sample more than once.

Our first estimate of the IGE between fathers and sons (Model 1, column 1 of Table 1a), based on a model in which we define fathers' lifetime earnings as the natural logarithm of average annual earnings from age 20 to 55 including years of zero earnings and do not exclude any fathers, is 0.299. While some of the years of zero earnings may be 'true' zeros, others may represent years in which earnings were not covered by Social Security. In particular, years in which all earnings are from military service or self-employment will appear as zeros in the administrative data. The IGE is not very sensitive to restricting the sample to fathers whose pre-1984 earnings are likely to be predominately covered by Social Security—in particular, non-veterans and non-government and non-self-employed workers. Dropping fathers who had ever served in the armed forces (as of 1984, when they were asked about their veteran status) reduces our sample size by almost half and reduces the estimated IGE in earnings slightly to 0.259 (see Model 2, column 1).[14] Restricting the sample of fathers to those not in the government or self-employed sectors yields an estimate of the IGE in earnings of 0.280 (Model 3, column 1).

The next 4 models of Table 1a show the sensitivity of the estimated IGE in earnings between fathers and sons to alternative definitions of fathers' lifetime earnings, definitions that eliminate years of zero earnings or remove observations with many years of zero earnings. These include defining fathers' lifetime earnings as the log of average earnings from the beginning of the first five consecutive years of positive earnings to age 55 (Model 4), as the log of average positive earnings from age 20 to 55 (Model 5), restricting the sample to fathers with 16 or more years of positive earnings (Model 6), and restricting the sample to fathers with positive earnings at every age from age 25 to 55 (Model 7).

Averaging over years beginning with the first five consecutive years of positive earnings leads to higher estimates of the IGE, 0.498 (see Model 4, column 1), as does averaging only over years of positive earnings when constructing fathers' lifetime earnings—0.507 (see Model 5, column 1). Restricting the sample of fathers to those with 16 or more years of positive earnings also leads to higher estimates: 0.482

[14]Roughly 48 percent of our fathers reported being a veteran. In the 1984 March Current Population Survey, 46 percent of men born between 1931 and 1949 (the same cohort as our sample of fathers) reported being a veteran. Roughly half of those veterans served during the time of the Vietnam conflict and roughly one-quarter served during the time of the Korean War.

(Model 6, column 1). Further restricting the sample of fathers to those with positive earnings at every age between age 25 and 55 yields a very high estimates of the IGE: 0.632 (Model 7, column 1).

Our first set of estimates of the IRA in earnings between fathers and sons, presented in column 2 of Table 1a, again differ from one another in terms of sample or in the definition of fathers' lifetime earnings. Unlike the estimates of the IGE in earnings, our estimates of the IRA in earnings between fathers and their sons are relatively invariant to changes in sample or in definition of lifetime earnings, ranging only from 0.292 to 0.395, and are centered at roughly 0.3.

In our first estimate of the IRA in earnings between fathers and sons (Model 1, column 2 of Table 1a), we define fathers' lifetime earnings as the average annual earnings from age 20 to 55 including years of zero earnings and do not exclude any fathers; this estimate is 0.292. As with the estimates of the IGE, this estimate is not sensitive to restricting the sample to non-veterans or to non-government and non-self-employed workers, restrictions which yield estimates of 0.314 and 0.320, respectively (see Models 2 and 3 in column 2). Unlike the IGE estimates, however, the estimates of the IRA in earnings are not sensitive to alternative definitions of fathers' lifetime earnings that eliminate years of zero earnings or remove observations with many years of zero earnings. Averaging over years beginning with the first five-year span of positive earnings yields an estimate of the IRA of 0.313 (Model 4, column 2) and averaging only over years of positive earnings when constructing fathers' lifetime earnings yields an estimate of 0.323 (Model 5, column 2). Restricting the sample of fathers to those with 16 or more years of positive earnings yields an estimate of 0.293 (Model 6, column 2), while further restricting the sample of fathers to those with positive earnings at every age between ages 25 and 55 yields a relatively high estimate of 0.395 (Model 7, column 2).

The estimates of the IGE and IRA in earnings between fathers and daughters, reported in Table 1b, are consistently lower than the corresponding estimates between fathers and sons. Like the estimates for sons, the IGE estimates for daughters also vary considerably as we change the sample or the definition of lifetime earnings for fathers across models. The resulting estimates range from -0.041 to 0.269. Also like the estimates for sons, the estimates of the IRA in earnings between fathers and daughters are robust to sample changes and to changes in the definition of fathers' lifetime earnings, ranging only from 0.103 to 0.166.

The Effects of Life-Cycle and Attenuation Bias on Estimates of the IGE and IRA in Earnings

By using single-age measures of sons' and daughters' earnings at different ages, we are able to explore the effect of left-hand side life-cycle bias on both the estimates of the IGE in earnings and the estimates of the IRA in earnings.[15] These estimates are reported in Tables 2a and 2b. When earnings are measured at any age when the children are in their 30s, the estimates of both the IGE and of the IRA are close in magnitude to the corresponding results from Tables 1a and 1b (see Table 2a and 2b). By contrast, when earnings are measured when the children are in their 20s, both the estimates of the IGE and those of the IRA are biased substantially towards zero.[16]

Tables 3a and 3b explore the sensitivity of the estimates to the ages over which a 16-year average of fathers' earnings is constructed and used to measure lifetime earnings. In these models, the dependent variable is always children's log average earnings from age 35 to 37 while the independent variable is the log of 16 years of fathers' earnings averaged over various ages (25 to 40; 30 to 45; 35 to 50; and 40 to 55). These estimates of the IGE are biased towards zero (due to attenuation bias from using only 16-year averages of earnings) and vary with the age of the father, confirming that right-hand side life-cycle bias is also an issue, as one would expect from Haider and Solon (2006) and Bohlmark and Lindquist (2006).

Attenuation bias causes the estimates of the IGE to decrease as we average fathers' earnings over a smaller number of years (see Tables 4a and 4b). We again use children's log average earnings from age 35 to 37 as the dependent variable but vary the number of years over which we average fathers' earnings

[15]The use of single-year measures of earnings as a dependent variable does not induce any additional attenuation bias, despite the single year measures being noisier than a three-year average.

[16]The same pattern of life-cycle bias appears in estimates that eliminate years of zero earnings or that drop observations with many years of zero earnings. These estimates are available from the authors upon request.

Table 2a
Sensitivity of Estimates of the Intergenerational Elasticity and of the Intergenerational Rank Association between Sons' Earnings and Fathers' Lifetime Earnings to the Age at Which Sons' Earnings are Measured

Age of Son at which Earnings are Measured	(1) Intergenerational Elasticity (IGE)			(2) Intergenerational Rank Association (IRA)		
	Coef	SE	n	Coef	SE	n
Age 20	0.039	0.037	1,539	-0.021	0.027	1,539
Age 22	0.061	0.037	1,717	0.022	0.026	1,717
Age 24	0.167	0.034	1,712	0.146	0.025	1,712
Age 26	0.220	0.042	1,719	0.214	0.024	1,719
Age 28	0.239	0.039	1,728	0.262	0.024	1,728
Age 30	0.334	0.046	1,709	0.291	0.024	1,709
Age 32	0.324	0.044	1,681	0.286	0.025	1,681
Age 34	0.313	0.057	1,480	0.289	0.026	1,480
Age 36	0.247	0.055	1,166	0.283	0.030	1,166
Age 38	0.288	0.064	738	0.268	0.038	738
Sample	All fathers					
Definition of Father's Lifetime Earnings	Log of average earnings from age 20 to 55 including years of zero earnings					

Source: Authors' calculations from 1984 SIPP-SSA matched file.
Notes: The dependent variable is sons' log earnings at the ages indicated in column (1) and is sons' position in distribution of earnings at the age indicated.
Fathers' lifetime earnings are defined as indicated.
Standard errors are adjusted for the fact that some fathers appear in the sample more than once.

Table 2b
Sensitivity of Estimates of the Intergenerational Elasticity and of the Intergenerational Rank Association between Daughters' Earnings and Fathers' Lifetime Earnings to the Age at Which Daughters' Earnings are Measured

Age of Daughter at which Earnings are Measured	(1) Intergenerational Elasticity (IGE)			(2) Intergenerational Rank Association (IRA)		
	Coef	SE	n	Coef	SE	n
Age 20	0.086	0.052	1,314	0.010	0.029	1,314
Age 22	0.114	0.050	1,444	0.055	0.027	1,444
Age 24	0.226	0.048	1,427	0.213	0.027	1,427
Age 26	0.314	0.060	1,439	0.231	0.028	1,439
Age 28	0.264	0.047	1,401	0.211	0.028	1,401
Age 30	0.276	0.055	1,385	0.206	0.028	1,385
Age 32	0.221	0.050	1,353	0.180	0.029	1,353
Age 34	0.229	0.063	1,159	0.156	0.031	1,159
Age 36	0.150	0.067	924	0.130	0.034	924
Age 38	0.192	0.101	571	0.133	0.045	571
Sample	All fathers					
Definition of Father's Lifetime Earnings	Log of average earnings from age 20 to 55 including years of zero earnings					

Source: Authors' calculations from 1984 SIPP-SSA matched file.
Notes: The dependent variable is daughters' log earnings at the ages indicated in column (1) and is daughters' position in distribution of earnings at the age indicated.
Fathers' lifetime earnings are defined as indicated.
Standard errors are adjusted for the fact that some fathers appear in the sample more than once.

Table 3a
Sensitivity of Estimates of the Intergenerational Elasticity and of the Intergenerational Rank Association between Sons' Earnings and Fathers' Lifetime Earnings to the Age Span Over Which Fathers' Earnings are Measured

16-Year Span of Ages over which Fathers' Earnings are Measured	(1) Intergenerational Elasticity (IGE)			(2) Intergenerational Rank Association (IRA)		
	Coef	SE	n	Coef	SE	n
Age 25 to 40	0.117	0.041	983	0.240	0.034	983
Age 30 to 45	0.249	0.051	986	0.280	0.033	986
Age 35 to 50	0.247	0.041	993	0.289	0.033	993
Age 40 to 55	0.272	0.055	985	0.287	0.034	985
Sample	All fathers					
Definition of Fathers' Lifetime Earnings	Log of average earnings from over the years indicated including years of zero earnings					

Source: Authors' calculations from 1984 SIPP-SSA matched file.
Notes: The dependent variable is sons' log earnings from age 35 and 37 in column (1) and sons' position in the earnings distribution at age 35 to 37.
Fathers' lifetime earnings are defined as indicated.
Standard errors are adjusted for the fact that some fathers appear in the sample more than once.

Table 3b
Sensitivity of Estimates of the Intergenerational Elasticity and of the Intergenerational Rank Association between Daughters' Earnings and Fathers' Lifetime Earnings to the Age at Which Child Earnings are Measured

16-Year Span of Ages over which Fathers' Earnings are Measured	(1) Intergenerational Elasticity (IGE)			(2) Intergenerational Rank Association (IRA)		
	Coef	SE	n	Coef	SE	n
Age 25 to 40	0.075	0.058	796	0.087	0.038	796
Age 30 to 45	0.114	0.065	804	0.098	0.039	804
Age 35 to 50	0.124	0.056	814	0.132	0.038	814
Age 40 to 55	0.152	0.051	816	0.161	0.037	816
Sample	All fathers					
Definition of Fathers' Lifetime Earnings	Log of average earnings from over the years indicated including years of zero earnings					

Source: Authors' calculations from 1984 SIPP-SSA matched file.
Notes: The dependent variable is daughters' log earnings from age 35 and 37 in column (1) and daughters' position in the earnings distribution at age 35 to 37.
Fathers' lifetime earnings are defined as indicated.
Standard errors are adjusted for the fact that some fathers appear in the sample more than once.

Table 4a
Evidence of Attenuation Bias in Estimates of the Intergenerational Elasticity and of the Intergenerational Rank Association between Sons' Earnings and Fathers' Lifetime Earnings

Span of Ages over which Fathers' Earnings are Measured	(1) Intergenerational Elasticity (IGE)			(2) Intergenerational Rank Association (IRA)		
	Coef	SE	n	Coef	SE	n
Age 20 to 55 (36 Years)	0.299	0.069	1,000	0.292	0.033	1,000
Age 25 to 55 (31 Years)	0.294	0.065	1,000	0.300	0.033	1,000
Age 30 to 50 (21 Years)	0.252	0.054	998	0.283	0.033	998
Age 30 to 40 (11 Years)	0.197	0.050	970	0.269	0.034	970
Age 35 to 40 (6 Years)	0.173	0.035	955	0.274	0.034	955
Age 35 to 37 (3 Years)	0.197	0.046	925	0.291	0.035	925
Sample	All fathers					
Definition of Fathers' Lifetime Earnings	Log of average earnings over the years indicated including years of zero earnings					

Source: Authors' calculations from 1984 SIPP-SSA matched file.
Notes: The dependent variable is sons' log earnings from age 35 to 37 in column (1) and sons' position in distribution of earnings averaged from age 35 to 37 in column (2).
Fathers' lifetime earnings are defined as indicated.
Standard errors are adjusted for the fact that some fathers appear in the sample more than once.

Table 4b
Evidence of Attenuation Bias in Estimates of the Intergenerational Elasticity and of the Intergenerational Rank Association between Daughters' Earnings and Fathers' Lifetime Earnings

Span of Ages over which Fathers' Earnings are Measured	(1) Intergenerational Elasticity (IGE)			(2) Intergenerational Rank Association (IRA)		
	Coef	SE	n	Coef	SE	n
Age 20 to 55 (36 Years)	0.177	0.073	820	0.120	0.037	820
Age 25 to 55 (31 Years)	0.177	0.069	820	0.125	0.037	820
Age 30 to 50 (21 Years)	0.133	0.061	816	0.121	0.038	816
Age 30 to 40 (11 Years)	0.123	0.065	783	0.100	0.039	783
Age 35 to 40 (6 Years)	0.105	0.057	774	0.099	0.040	774
Age 35 to 37 (3 Years)	0.126	0.073	760	0.100	0.040	760
Sample	All fathers					
Definition of Fathers' Lifetime Earnings	Log of average earnings over the years indicated including years of zero earnings					

Source: Authors' calculations from 1984 SIPP-SSA matched file.
Notes: The dependent variable is daughters' log earnings from age 35 to 37 in column (1) and daughters' position in distribution of earnings averaged from age 35 to 37 in column (2).
Fathers' lifetime earnings are defined as indicated.
Standard errors are adjusted for the fact that some fathers appear in the sample more than once.

(20 to 55; 25 to 55; 30 to 50; 30 to 40; 35 to 40; and 35 to 37). Surprisingly, estimates of the IRA in earnings are robust to this choice.

To assess how our estimates of the IGE are sensitive to choice of specification, data, and other choices, we conduct a large number of sensitivity analyses for the estimates of the IGE between fathers and sons (see Table 5).[17] As discussed above, previous studies have typically estimated a "calendar year" specification (Equation 1) as opposed to our "lifetime" specification (Equation 2), which controls for both life-cycle and attenuation biases. In an important study, Mazumder (2005) estimated a calendar year specification using similar data to ours—the 1984 SIPP matched to SER data.[18]

Our analysis shows that the estimates of the IGE are sensitive to choice of specification.[19] The particular calendar year specification we estimate follows Mazumder (2005) and is:

$$y_{it} = \alpha + \beta x_{it} + f\left(\gamma, Age_i\right) + \varepsilon_{it} \tag{4}$$

where:

y_{it} is the log of average child's earnings over the 1995 to 1998 period;

x_{it} is the log of average father's earnings over the 1970 to 1985 period; and

f(.) is a high-order polynomial function in father's and child's age.

Moving from the lifetime specification (based on Equation 2) of 0.299 (in row 1, column 1 of Table 5) to a calendar year specification (Equation 4) on the same data (row 2) reduces the estimate of the IGE to 0.275. We observe a similar reduction in the estimate when we eliminate years of zero earnings (see column 2) or restricting the sample of fathers to those with at least 16 years of positive earnings (see column 3).

---

[17]Results of the IGE sensitivity analyses for fathers and daughters and the IRA sensitivity analyses are available upon request.

[18]This study also slightly differs from ours in a number of other analytic and data construction choices, the effects of which we have explored and the results of which are available upon request.

[19]Our analysis (not shown) shows that the choice of age ranges of children and imputation method make little or no difference to the estimates.

Table 5
Sensitivity of IGE Estimates for Fathers and Sons to Choice of Specification and Sample

Specification	Measure of Child's Earnings	Measure of Father's Earnings	(1) Coef	(1) SE	(1) n	(2) Coef	(2) SE	(2) n	(3) Coef	(3) SE	(3) n
Lifetime	Age 35 to 37	Age 20 to 55	0.299	0.069	1,000	0.507	0.082	1,000	0.482	0.074	949
Calendar	1995 to 1998	1970 to 1985	0.275	0.040	1,793	0.411	0.054	1,793	0.327	0.053	1,709
Sample				All fathers			All fathers			Fathers with 16 or more years of earnings	
Definition of Father's Lifetime Earnings				Log of average earnings including years of zero earnings			Log of average of years of positive earnings			Log of average earnings including years of zero earnings	

Source: Authors' calculations from 1984 SIPP-SSA matched file.
Notes: Sons' earnings are the log of average earnings over the ages or years indicated.
Fathers' lifetime earnings are defined as indicated.
Standard errors are adjusted for the fact that some fathers appear in the sample more than once.

In summary, both the estimates of the IGE in earnings and of the IRA in earnings require measuring earnings of children when they are in their 30s. Estimates based on earnings of children in their 20s appear to be substantially downward biased. Estimates of the IGE are sensitive to the age spans over which father's earnings are averaged, while the estimates of the IRA are less sensitive. Estimates of the IGE are also sensitive to the number of years of fathers' earnings used while estimates of the IRA, remarkably, are not sensitive. This finding may be of particular use to researchers who lack access to the long earnings histories available to us from which we construct fathers' lifetime earnings.

VI. NONPARAMETRIC MEASURES OF INTERGENERATIONAL MOBILITY

The estimates of the IGE and of the IRA in earnings presented above provide single-number measures of the degree of intergenerational mobility. In this section, we discuss and present nonparametric estimates of intergenerational mobility from which one can observe a complex and nonlinear relationship between fathers' and children's earnings.

Transition Matrices

A commonly used method of expressing mobility is a transition matrix. This matrix shows the probability of a child's being in a given quintile of the earnings distribution conditional upon his or her father's position in the lifetime earnings distribution. Table 6 presents a transition matrix by quintile for children and fathers. The table shows a fair degree of earnings mobility across generations. For example, only 29 percent of sons and 25 percent of daughters whose fathers are in the lowest quintile of the lifetime earnings distribution are themselves in the lowest quintile, and 10 percent of those sons and 17 percent of those daughters transition into the highest quintile. Forty-one percent of sons and 31 percent of daughters with fathers in the highest quintile of the lifetime earnings distribution are themselves in the highest quintile. Roughly 15 percent to 23 percent of those children transition into the lowest quintile.

These estimates are remarkably similar to those presented by Peters (1992) and by Zimmerman (1992) using the NLS. As we note above, we find it interesting that these two studies found widely

Table 6
Transition Matrices between Children's Earnings and Fathers' Lifetime Earnings

		Son's Earnings, by Quintile				
		Bottom	2nd	3rd	4th	Top
Father's Earnings, by Quintile	Bottom	28.9	26.9	18.8	15.7	9.6
	2nd	23.2	25.3	21.7	15.5	14.4
	3rd	18.7	20.2	23.7	23.2	14.1
	4th	15.6	14.2	22.7	25.8	21.8
	Top	14.0	14.0	12.4	18.8	40.9
		Daughter's Earnings, by Quintile				
		Bottom	2nd	3rd	4th	Top
Father's Earnings, by Quintile	Bottom	24.7	21.6	21.6	15.4	16.7
	2nd	19.5	20.8	24.7	18.2	16.9
	3rd	17.7	24.1	22.4	21.8	14.1
	4th	15.3	17.2	21.0	25.5	21.0
	Top	22.6	16.4	11.3	19.2	30.5

Source: Authors' calculations from the 1984 SIPP-SSA matched file.
Notes: Children's earnings are the log of average earnings from age 35 and 37.
Fathers' lifetime earnings are the log of average earnings from age 20 to 55.
Years with zero earnings are included in fathers' lifetime earnings and in multi-age averages of children's earnings.

differing estimates of the IGE in earnings for fathers and sons—0.14 in Peters (1992) and 0.54 in Zimmerman (1992)—yet similar estimates of the transition matrix. This robustness of these estimates across studies, like the robustness of our estimates of the IRA in earnings across samples and specifications, provides further evidence that the methodological and sample choice issues that have led to the wide range of estimates of the IGE in earnings may be less of an issue in estimates based on relative earnings position.

Nonlinear Estimates of the IRA

The IRA models we estimate impose a linear relationship between the position of children in the earnings distribution and the position of fathers in the lifetime earnings distribution and thus, like the IGE, yield a single number for the degree of intergenerational mobility. In order to allow for a potentially nonlinear and highly flexible relationship between children's and fathers' positions in their respective earnings distributions, we calculate a simple yet highly flexible nonparametric estimate of this relationship.

To do this, we create moving blocks (of five percentiles) of fathers ordered by their percentile in the lifetime earnings distribution. For example, the first block includes all fathers in the first through fifth percentiles of lifetime earnings; the second block includes fathers in the second through sixth percentiles; and the 96th block includes fathers in the 96th through 100th percentiles. For each block, we calculate the percentile in the overall children's earnings distribution at which the 20th percentile, median, and 80th percentile of earnings of the children in that block fall (where earnings of children once again are measured as the average between ages 35 and 37). We do this calculation separately for sons and daughters.

Figures 1 and 2 display this nonparametric relationship between children's relative positions in the earnings distribution (measured at ages 35 through 37) and fathers' relative positions in the lifetime earnings distribution. The solid line displays the median percentile of the children's positions in the earnings distribution. Roughly speaking, a child has a 50 percent probability of having earnings below this line given his or her father's position in the lifetime earnings distribution. The two dashed lines display the equivalent 20th and 80th percentiles.

There is a substantial amount of mobility between fathers and sons in the middle of the earnings distribution (see Figure 1). Sons whose fathers are between the 10th and 80th percentile of the lifetime earnings distribution have a roughly equal distribution of earnings. For example, sons whose fathers have about median lifetime earnings have roughly an 80 percent chance of being below the 80th percentile of earnings, a 50 percent chance of being below median earnings, and a 20 percent chance of being below the 20th percentile for earnings. Sons whose fathers have about the 10th percentile of lifetime earnings have roughly an 80 percent chance of being below the 75th percentile, a 50 percent change of being below the 45th percentile, and a 20 percent change of being below the 20th percentile. Similarly, for sons whose fathers have about the 80th percentile of lifetime earnings, these percentiles are the 80th, the 60th, and the 30th. Thus, while there is a positive association between fathers' and sons' positions in the earnings distribution, it is very small in the "broad" middle of the lifetime earnings distribution. Between the 10th and 80th percentiles, the linear relationship between fathers' and sons' earnings percentiles is about 0.2.

This large amount of mobility for the bulk of fathers is in striking contrast to the amount at either tail of the distribution of fathers' lifetime earnings. Both above the 80th percentile and below the 10th, the association between fathers' position and sons' is very high. That is, there is much less intergenerational mobility in earnings at the top (above the 80th) and at the extreme bottom (below the 10th) of the earnings distribution.

Figure 2 displays the equivalent relationship between fathers and daughters. Like the "broad middle" in the case of sons, there is only a weak and positive relationship between fathers' and daughters' position in the earnings distribution. However in contrast to sons, this weak association is evident at the tails of the distribution as well. In fact, for daughters whose fathers are above the 90th percentile in lifetime earnings, this association appears to be negative. This negative association at the top of the

Figure 1
Son's Position in the Earnings Distribution by Father's Position in the Earnings Distribution
(Moving Block of 5 Positions)

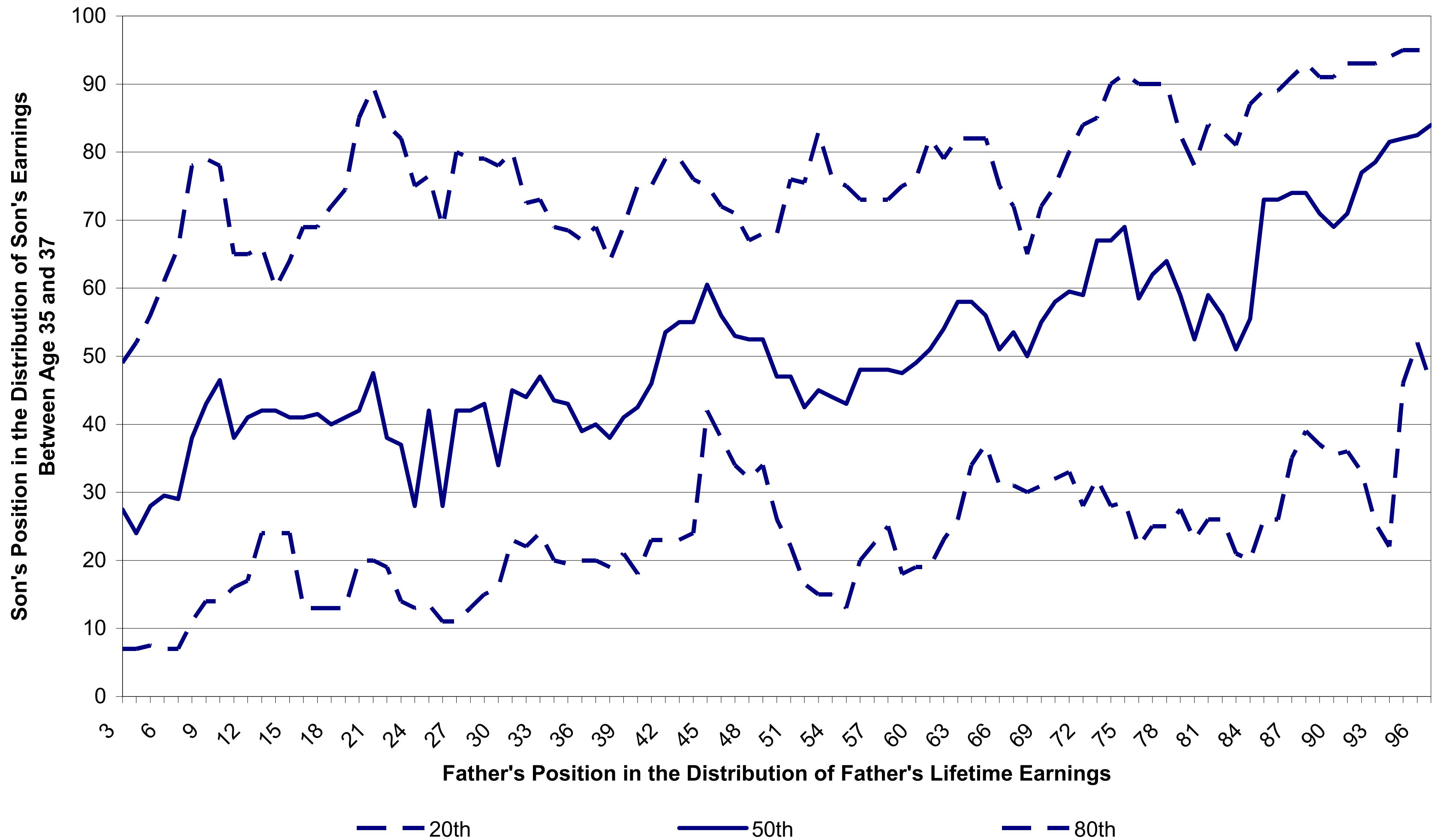
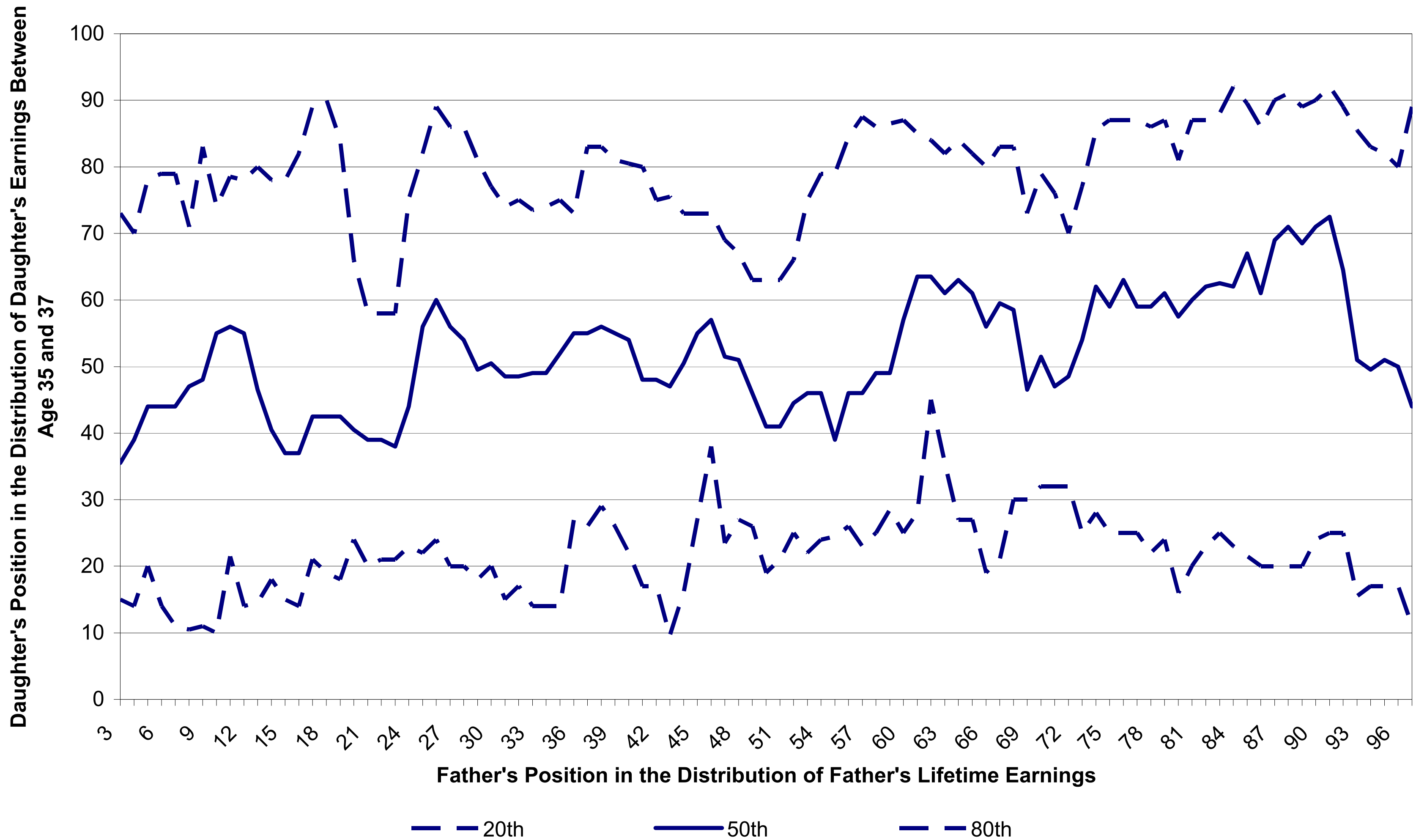


Figure 2
Daughter's Position in the Earnings Distribution by Father's Position in the Earnings Distribution
(Moving Block of 5 Positions)



distribution might be explained by these daughters having, or expecting to have, high-earnings husbands.[20]

VII. CONCLUSION

We show that estimates of the IGE in earnings are highly sensitive to choice of sample, definition of fathers' lifetime earnings, and specification. Depending upon small changes in sample or in the definition of lifetime earnings, for example, we estimate the IGE in earnings between sons and fathers to be as low as 0.26 or as high as 0.63, estimates which, alternatively, imply either a high or a low degree of intergenerational mobility.

By contrast, estimates of the association between fathers' and children's relative positions in the earnings distribution (the IRA in earnings) demonstrate that there is substantial intergenerational mobility and, moreover, these estimates are not as sensitive to the choice of specification, sample, or inclusion of years of zero earnings as are the IGE estimates. These estimates suggest that a 10 percentage point increase in a father's position in the lifetime earnings distribution is associated with sons' being, on average, roughly 3 percentiles higher and daughters' being, on average, roughly 1 percentile higher in their respective earnings distributions.

We are able to provide this wide array of estimates of the extent of intergenerational mobility through the use of high-quality and nearly career-long SSA earnings histories of a large sample of fathers. We measure sons' and daughters' earnings around an age—36—that is likely to be a good proxy for lifetime earnings according to recent estimates (Haider and Solon, 2006; Bohlmark and Lindquist, 2006) and confirmed by our own analysis.

Nonparametric estimates of the association between fathers' and children's relative positions in the earnings distribution show that there is substantial mobility in the broad middle of the distribution (between the 10^{th} and 80^{th} percentiles of the distribution of fathers' lifetime earnings) for sons and over

---

[20] Lam and Schoeni (1993) found evidence of an association between the earnings of sons-in-law and fathers using data from Brazil.

the entire distribution for daughters. For sons whose fathers are at the tails of the lifetime earnings distribution (above the 80th or below the 10th), there is very little mobility. When a father has very high or very low lifetime earnings, his position is a strong determinant of his son's earnings.

Given the wide variation in the estimates of IGE, it is difficult to draw strong conclusions about the extent of intergenerational earnings mobility between fathers and their children based on these estimates. By contrast, the estimates of the IRA in earnings are robust across specifications and samples, and thereby provide more conclusive evidence on the extent of intergenerational mobility. A few studies (e.g., Corak, 2006; and Jäntti, 2008) provide cross-national estimates of the IGE and have found the U.S. to be relatively immobile. Our analysis suggests that these cross-national comparisons may be very difficult to make given the high level of sensitivity of the IGE to small changes in specification and sample, and that future cross-national work should consider comparing estimates of the IRA in earnings.

Appendix

We use a sample of earnings for men from the DER to impute earnings for men from the SER, which are potentially subject to top-coding. Table A2 reports the Social Security taxable maximum in nominal and real dollars and the percentage of fathers with top-coded earnings data by year, from 1951 to 1983. Between 1951 and 1964, the top-codes were rather low, roughly $22,000 to $26,000 in 2005 dollars. Our sample of fathers, who were born between 1931 and 1949, are relatively young in those years so that, at least initially, relatively few of them have annual earnings above the top-code. However, by 1965, almost half of our sample of fathers has top-coded earnings. Seventy-five percent of our sample have at least one year of earnings that is above the top-code while 42 percent have 10 or more years of top-coded data. After 1983 we use earnings from the DER, which are not top-coded.

In order to impute top-coded earnings, we use propensity score methods to match top-coded observations from the SER sample to uncensored observations from the DER sample. We use the rich panel data on earnings available in both surveys as our matching variables. In particular, we estimate:

$$P\left(y_{a,t} \geq T_t\right) = F\left(\alpha + \sum_{j\in(-5,-1)(1,5)} \beta_j y^*_{(a+j),(t+j)} + \sum_{j\in(-5,-1)(1,5)} \gamma_j I\left(y_{(a+j),(t+j)} \geq T_{t+j}\right)\right)$$

where

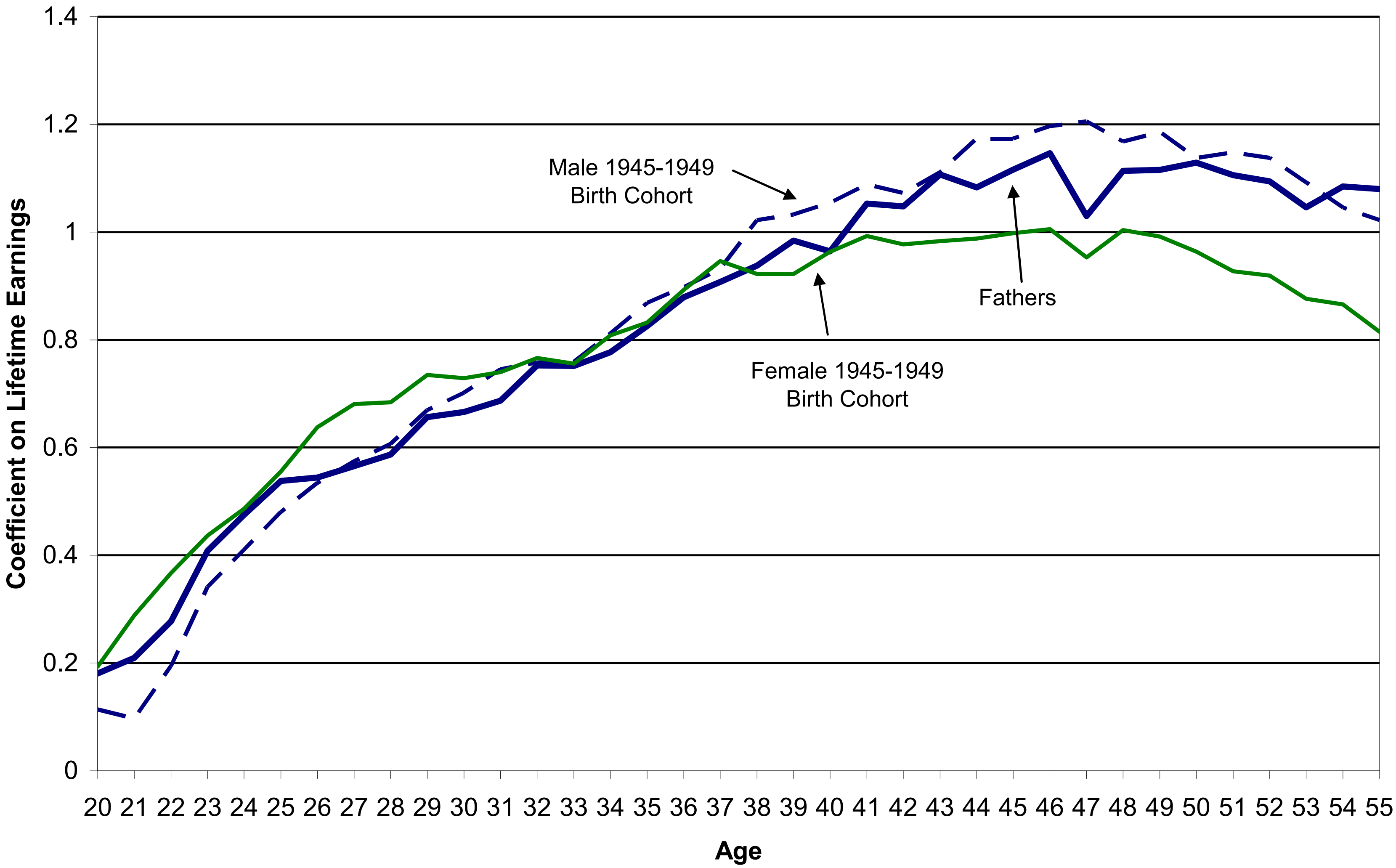
$y_{a,t}$ is the position in the earnings distribution of uncensored earnings at age *a* in year *t*;

T_t is the position in the earnings distribution of the Social Security taxable maximum in year *t*;

$y^*_{a,t}$ is equal to T_t if censored and to $y_{a,t}$ is uncensored.

We measure these positions as half centiles in the current-year earnings distribution. We use this model to construct propensity scores for both the SER and DER samples and use propensity score methods to match each top-coded observation from the SER to 20 uncensored observations from the DER using a nearest neighbor method. The imputed value is the average of the values from the matched sample. The imputed values are then calculated using the distribution of earnings from the Current Population Survey for that year (we use the 1963 CPS for the years 1951 through 1962).

Appendix Figure 1
Estimate of Life-Cycle Bias: Annual Earnings on Lifetime Earnings



Appendix Table 1
Estimates of the Degree of Life-Cycle Bias:
Coefficients from Regressions of Multi-Age Averages of Annual Earnings on Lifetime Earnings

	Fathers Sample	1945–1949 Cohort	
		Men	Women
Age 35 to 37	0.97	1.02	1.02
Age 34 to 38	1.01	1.04	1.04
Age 33 to 39	1.04	1.05	1.05
Age 32 to 40	1.03	1.06	1.05

Source: Authors' calculations from the SIPP-SSA matched files.
Notes: Our sample of fathers were born between 1931 and 1949 and had at least one child born between 1963 and 1972 who was living with him at the time of the first 1984 SIPP interview (which took place at some point between October 1983 and January 1984).

Appendix Table 2
Percentage of Fathers with Top-Coded SER Earnings by Year, 1951 to 1983

Year	Social Security Taxable Maximum: In Nominal Dollars	Social Security Taxable Maximum: In Real (2005) Dollars	Percent of Fathers w/ Top-Coded Earnings in that Year	Percent of Fathers w/ Age ≥ 20 and ≤ 55 in that Year	Percent of Fathers w/ Top-Coded Earnings Conditional on Age ≥ 20 and ≤ 55 in that Year
1951	$3,600	$22,857	0.0	3.6	0.0
1952	3,600	22,460	0.1	8.1	1.7
1953	3,600	22,305	0.6	13.2	4.3
1954	3,600	22,152	0.8	19.1	4.4
1955	4,200	25,933	1.5	24.5	6.1
1956	4,200	25,495	3.4	30.1	11.3
1957	4,200	24,742	5.2	36.5	14.2
1958	4,200	24,033	8.0	43.3	18.4
1959	4,800	27,293	8.9	49.8	17.8
1960	4,800	26,784	12.5	55.7	22.5
1961	4,800	26,537	16.1	62.0	25.9
1962	4,800	26,294	21.6	69.5	31.0
1963	4,800	25,899	27.3	76.3	35.7
1964	4,800	25,592	34.0	81.6	41.6
1965	4,800	25,218	40.8	85.8	47.5
1966	6,600	33,690	30.9	90.6	34.1
1967	6,600	32,669	39.0	94.6	41.2
1968	7,800	37,175	35.8	97.4	36.8
1969	7,800	35,571	44.2	100.0	44.2
1970	7,800	33,934	50.1	100.0	50.1
1971	7,800	32,517	56.9	100.0	56.9
1972	9,000	36,421	53.9	100.0	53.9
1973	10,800	41,113	47.7	100.0	47.7
1974	13,200	45,698	38.1	100.0	38.1
1975	14,100	45,079	37.2	100.0	37.2
1976	15,300	46,281	38.6	100.0	38.6
1977	16,500	46,910	40.5	100.0	40.5
1978	17,700	48,247	42.4	100.0	42.4
1979	22,900	57,059	30.8	100.0	30.8
1980	25,900	58,117	26.9	100.0	26.9
1981	29,700	60,921	25.0	100.0	25.0
1982	32,400	62,747	22.0	100.0	22.0
1983	35,700	66,380	19.8	100.0	19.8

Source: U.S. Social Security Administration (www.ssa.gov) and authors' calculations from the SIPP-SSA matched files.

Memo: Percent of sample with:

1 or more years of top-coded data	74.9
5 or more years of top-coded data	57.0
10 or more years of top-coded data	42.3

Notes: Social Security taxable maximum inflated to 2005 dollars using CPI-U-RS
Fathers appearing in the data more than once (because they have multiple children) are included here multiple times. Results are similar if they are only included once. Sample Size: 3,521.

References

Becker, Gary S., and Nigel Tomes. 1986. "Human Capital and the Rise and Fall of Families." *Journal of Labor Economics* 4(July): S1–S39.

Beller, Emily, and Michael Hout. 2006. "Intergenerational Social Mobility: The United States in Comparative Perspective." *The Future of Children* 16(2, Fall): 19–36.

Blau, Peter, and Otis Dudley Duncan. 1967. *The American Occupational Structure*. New York: Wiley.

Bjorklund, Anders. 1993. "A Comparison between Actual Distributions of Annual and Lifetime Income: Sweden 1951–89." *Review of Income and Wealth* 39(4, December): 377–86.

Bohlmark, Anders, and Matthew J. Lindquist. 2006. "Life-Cycle Variations in the Association between Current and Lifetime Income: Country, Gender and Cohort Differences." *Journal of Labor Economics* 24(4, October): 879–96.

Bound, John, and Alan B. Krueger. 1991. "The Extent of Measurement Error in Longitudinal Earnings Data: Do Two Wrongs Make a Right?" *Journal of Labor Economics* 9(1, January): 1–24.

Chadwick, Laura and Gary Solon. 2002. "Intergenerational Income Mobility Among Daughters." *American Economic Review* 92(1, March): 335–44.

Corak, Miles. 2006. "Do Poor Children Become Poor Adults? Lessons from a Cross Country Comparison of Generational Earnings Mobility." *Research on Economic Inequality* 13(1, January): 143–88.

Corak, Miles, and Andrew Heisz. 1999. "The Intergenerational Earnings and Income Mobility of Canadian Men: Evidence from Longitudinal Income Tax Data." *Journal of Human Resources*, 34(3, Summer): 504–33.

Couch, Kenneth A., and Dunn, Thomas A. 1997. "Intergenerational Correlations in Labor Market Status: A Comparison of the United States and Germany." *Journal of Human Resources*, 32(1, Winter): 210–32.

Couch, Kenneth A., and Dean R. Lillard. 1998. "Sample Selection Rules and the Intergenerational Correlation of Earnings." *Labour Economics* 5(3, September): 313–29.

Eide, Eric R., and Mark H. Showalter. 1999. "Factors Affecting the Transmission of Earnings across Generations: A Quantile Regression Approach." *Journal of Human Resources* 34(2, Spring): 253–67.

Grawe, Nathan D. (Forthcoming). "Lifecycle Bias in Estimates of Intergenerational Earnings Persistence." *Labour Economics*.

Haider, Steven J., and Gary Solon. 2006. "Life-Cycle Variation in the Association between Current and Lifetime Earnings." *American Economic Review* 96(4, October): 1308–20.

Isaacs, Julia. 2008. *Economic Mobility of Families Across Generations*. Report of the Economic Mobility Project. Washington, DC: Brookings Institution.

Jäntti, Markus. 2008. “Mobility in the U.S. and in Comparative Perspective.” Working paper, Åbo, Finland: Åbo Akademi University.

Lam, David, and Robert F. Schoeni. 1993. “Effects of Family Background on Earnings and Returns to Schooling: Evidence from Brazil.” *Journal of Political Economy* 101(4, August): 710–40.

Peters, H. Elizabeth. 1992. “Patterns of Intergenerational Mobility in Income and Earnings.” *Review of Economics and Statistics* 74(3, August): 456–66.

Mazumder, Bhashkar. 2005. “Fortunate Sons: New Estimates of Intergenerational Mobility in the U.S. Using Social Security Earnings Data.” *Review of Economics and Statistics* 87(2, May): 235–55.

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

Solon, Gary. “Comments on ‘Sample Selection Rules and the Intergenerational Correlation of Earnings: A Comment on Solon and Zimmerman’ by Couch and Lillard.” Unpublished paper, University of Michigan.

Solon, Gary. 1999. “Intergenerational Mobility in the Labor Market.” In *Handbook of Labor Economics*, edited by Orley C. Ashenfelter and David Card, Vol. 3A. Amsterdam: North-Holland, 1761–800.

Vogel, Thorsten. 2007. “Reassessing Intergenerational Mobility in Germany and the United States: The Impact of Differences in Lifecycle Earnings Patterns.” Unpublished paper, Humboldt University.

Zimmerman, David J. 1992. “Regression Toward Mediocrity in Economic Stature.” *American Economic Review* 82(3, June): 409–29.