Paper Prepared fro the 28th General Conference of
The International Association for Research in Income and Wealth
Cork, Ireland, August 22 - 28, 2004

LIFE CYCLE BIAS IN THE ESTIMATION OF
INTERGENERATIONAL EARNINGS PERSISTENCE

NATHAN D. GRAWE

For additional information please contact:

Nathan D. Grawe
Department of Economics
Carleton College
Northfield, MN 55057
USA
Email: ngrawe@carleton.edu
Fax: (507) 646-4044
Telephone: (507) 646-5239

This paper is posted on the following websites: **http://www.iariw.org**
http://www.econ.nyu.edu/iariw
http://www.cso.ie/iariwhome.html

Abstract

Using data from the National Longitudinal Survey, the German Socio-Economic Panel, the Canadian Intergenerational Income Data, and the Panel Study of Income Dynamics; this study finds a strong association between estimated intergenerational earnings elasticities and the age at which fathers and sons are observed. In all four data sets the age-elasticity relationship (which is positive for sons and negative for fathers) is especially strong among fathers; estimates are cut in half as the fathers' ages at observation increase by fewer than fifteen years. These effects are consistent with either increasing transitory earnings variance (and so greater attenuation bias) or increasing earnings variance over the life cycle predicted by the human capital investment models of Mincer and Ben-Porath. Furthermore, an examination of published estimates of intergenerational earnings elasticities shows that controls for the average age of fathers explain about one-third of the variance among estimates. These results impact our interpretation of empirical work which attempts to differentiate between the importance of parent income when children are young as opposed to when children are older, work which has been used to draw conclusions about credit constraints and early childhood education programs.

I. Introduction

With recent data set developments, economists have been able to estimate the degree of intergenerational earnings persistence using data from multiple sources and multiple countries.[1] However, an examination of the resulting estimates paints a confusing picture with wide variation even among studies using data from the same source. For example, compare Solon's (1992) persistence estimate of 0.41 to Couch and Dunn's (1997) estimate of 0.13. Both studies use Panel Study of Income Dynamics (PSID) data. While we might not expect to find complete agreement among empirical studies, the immense differences found in this literature undermine efforts to draw conclusions from the data. Which of the disparate estimates from the PSID best represents mobility in the United States? Can we effectively compare mobility between countries? With no explanation as to why the observed variation exists, these questions are very difficult to answer.

Solon (1989) represents the first effort to harmonize empirical findings. Noting that earlier studies had relied on single-year observations of earnings that were doubtlessly affected to a greater or lesser degree by classical measurement error, Solon suspected that early estimates understated the degree of earnings persistence. By averaging multiple earnings reports and employing an instrumental variables (IV) approach, Solon (1992) corrected for this classical measurement error and showed that persistence estimates increased significantly as a result. To the extent that earlier works employed data with differing degrees of error, one might infer that some of the variation in estimates is explained by this factor.

But despite the important contribution of Solon's paper and widespread application of his

---

[1]Following the convention of the literature, the degree of earnings persistence is defined as the elasticity of sons' earnings with respect to fathers' earnings.

methodology, the variation in estimated intergenerational earnings persistence remains significant. Among fourteen published studies that correct for classical measurement error by averaging at least three earnings reports, the mean persistence estimate is 0.235 with standard deviation 0.118.[2] And this wide variation is found even within studies that use PSID data.

This paper explains a significant portion of this remaining variation by controlling for the age of fathers and sons at the point of earnings observation. Using data from the National Longitudinal Survey (NLS), the German Socio-Economic Panel (GSOEP), the Canadian Intergenerational Income Data (IID), and the PSID; the next section documents that intergenerational earnings persistence estimates are strongly sensitive to the age at which father's and son's earnings are measured. The third section shows that we might expect such effects using standard models of human capital accumulation. In the fourth section, these results are applied to the published intergenerational earnings persistence estimates showing that one-third of the variation among studies is explained by fathers' ages at observation. Finally the importance of these results to policy is explored. Several authors have found that parent earnings are a stronger predictor of child achievement when parents' earnings are measured early in life. The authors conclude that credit constraints significantly limit educational choice. The results of this paper suggest such that this association is easily explained without resorting to market imperfections.

II. Empirical Evidence of a Life Cycle Bias

Several existing studies provide some indications that intergenerational mobility estimates may a function of the ages of fathers and sons at observation. (Hereafter, this will be termed a 'life cycle bias.') For instance, Solon (1992) presents results for five years of father

---

[2]See Table 1 for a full listing of published estimates.

observation and Zimmerman's (1992) results cover four observations of fathers and seven observations of sons. Both studies show that earnings persistence estimates increase as sons age and decrease as fathers age. Clearly the limited time spans in these studies and the fact that both studies use only US data restrict our ability to draw strong conclusions. Reville (1995) explicitly studies age-dependency of intergenerational mobility estimates in the PSID. However, he only examines changes in the age of sons and the sample available to him at the time of writing restricted sons to relatively younger ages (data through 1990). Both of these limitations turn out to be significant. Recent developments in panel data sets permit a detailed examination of the life cycle bias. The data employed in this study are drawn from the National Longitudinal Survey (the NLS includes observations from nine years between 1966 and 1981), the German Socio-Economic Panel (1984-2001), the Canadian Intergenerational Income Data (1978-1998), and the Panel Study of Income Dynamics (1968-1993). See the appendix for detailed sample descriptions.

Each data set contains multiple observations for both sons and fathers.[3] Each earnings observation for the son is regressed in turn on each of the available earnings observations for the father including controls for both age and age-squared for both the father and the son.[4] For

---

[3]In order to avoid confounding effects of sample attrition due to the retirement of fathers, sample selection is limited by the age of fathers. For instance, in the examination of the PSID, fathers are no older than 46 in 1967 so that they are no older than 60 in 1981, the final year of observation. In other words, the graphs below follow a cohort of families across time.

[4]Obviously, single-year measures of earnings contain measurement error and so the levels of earnings persistence estimated in this section are lower than the true degree of persistence. However, in identifying the importance of a life cycle bias, we are interested in the *trend* in estimates over the life cycle – not the level of persistence itself. This trend is easier to identify when we have a large number of estimates from a wide range of ages. When the analysis is repeated using three-year averages of earnings, the same qualitative results obtain. But with one-third the number of independent persistence estimates, it is more difficult to determine whether the pattern constitutes a trend.

instance, in equation (1) sons' log earnings measured in 1993 are regressed on fathers' log earnings measured in 1987.

$$y_{s,93} = \alpha + \gamma_1 age_{s,93} + \gamma_2 age_{s,93}^2 + \gamma_3 age_{f,87} + \gamma_4 age_{f,87}^2 + \beta y_{f,87} \qquad (1)$$

β, the elasticity of son's earnings with respect to father's earnings, measures persistence in earnings across generations.

A. National Longitudinal Survey

Zimmerman (1992) reports earnings persistence estimates for multiple observations of both fathers and sons in the Original Cohort NLS. While Zimmerman's results are consistent with the presence of a large life cycle bias (estimated earnings persistence decreases/increases as fathers/sons age), the study restricts the sample to only include fathers and sons who are employed at least thirty hours per week and thirty weeks per year. I update the analysis using the more common restriction that respondents must report positive earnings to be included in the sample.

Just as in Zimmerman (1992), the persistence of earnings decreases as the year of father observation advances from 1965 to 1970 (see Figure 1). (All of the figures in this section that show age-dependence of persistence estimates are constructed in a similar fashion. Each year of father (son) observation can be matched with multiple years of son (father) observation. Earnings persistence estimates are computed for each possible father-son observation pair as described above. These estimates are represented by points in the figure. For instance, there are five NLS son observations that can be paired with each father observation. So, for each year of father observation, there are five estimates of earnings persistence. Since the standard errors are relatively large (approximately 0.05-0.10), I focus on the trend in the *average* of the persistence estimates which is represented by a solid line.) The reduction is quite large – more than fifty

percent as the age of the fathers increases five years.[5] However, cautious interpretation is required

due to the small number of observation points across the fathers' life cycles.

As sons' ages increase, Figure 2 reports a U-shape pattern in earnings persistence with

especially high persistence estimates when the observations are drawn at the latest point in the

sons' life cycles. However, once more the number of available observations is limiting. For

instance, are the results from 1980 an aberration or the beginning of an upward trend?

Fortunately, other data sources can resolve these problems.

B. German Socio-Economic Panel

While the GSOEP is too young to provide multiple measures of sons' earnings without

including very young sons or suffering from extremely small sample sizes, many observations of

fathers' earnings are available. Figure 3 plots persistence estimates resulting from regressions of

sons' log earnings in 2000 on fathers' log earnings in 1983 through 1997. The estimates of

earnings persistence decrease substantially as the year of father observation is increased. But the

relatively large standard errors (0.10 to 0.19) that result from the sample sizes in the GSOEP make

it difficult to draw strong conclusions from the German data. However, the rate at which

estimated earnings persistence in Germany diminishes across the fathers' life cycles is identical to

that found in the NLS – on average the elasticity diminishes by 0.0121 per year. Of course, it

would be preferable to have a data set with multiple measures of both fathers' and sons' earnings.

The Canadian IID and the American PSID contain just such data.

C. Intergenerational Income Data

With observations spanning fifteen years for fathers and eight years for sons, the Canadian

---

[5]In this and all subsequent figures showing the negative relationship between earnings
persistence estimates and the year of *father* observation, the relationship is statistically significant
($\alpha=0.02$).

IID allows a deeper examination of the life cycle bias. Figure 4 shows a substantial and sustained trend toward less earnings persistence as fathers age. (Since the IID sample size is so large, standard errors for each individual persistence estimate are very small – around 0.006 to 0.009.) In total, the estimated earnings persistence falls by more than one-third when the year of father observation is increased fifteen years. In absolute terms, this decrease is smaller than that found in the NLS and GSOEP: for each year that fathers age, on average the estimated elasticity diminishes by 0.004. Figure5 studies the effect of sons' ages on estimated earnings persistence. Increasing the age of son observation by eight years increases estimated earnings persistence by more than fifty percent or 0.004 per year.

D. Panel Study of Income Dynamics

By way of comparison, the PSID also contains multiple earnings observations for both fathers and sons over a relatively long time span. The PSID is of special importance since it is the source of most US studies of intergenerational earnings persistence. Figure 6 plots the resulting earnings persistence estimates as the period of observation for fathers is varied from 1967 to 1981. (The standard errors are very similar to those found in the NLS – around 0.05-0.10.) A clear downward trend is evident, a decrease of approximately fifty percent over fourteen years. The absolute rate of this trend is very similar to that in the NLS and GSOEP: on average estimated persistence diminishes by 0.0138 per year that fathers age. Persistence patterns as sons age provide a second opportunity to identify a life cycle bias. Figure 7 plots the estimates of earnings persistence against year of son observation for each period of father observation. In the time period 1983-1990 which is common to this study and Reville (1995) the pattern is very much like that reported by Reville. The persistence estimates vary significantly across years of son observation, but there is no evidence of trend. The additional years included in this paper

(1991-1992) show that earnings persistence may be trending upward as sons age.

In summary, in all four samples studied clear evidence was found for a strong negative relationship between the age of fathers at observation and estimated intergenerational earnings persistence. Moreover, the magnitude of the effect was very similar in all four countries. Persistence estimates decreased by around 0.0125 per year that fathers aged in the NLS, GSOEP, and PSID. The relationship was weaker in Canada where persistence estimates fell by around 0.004 per year that fathers aged. The relationship between the age of sons at observation and earnings persistence was less clear due in part to data limitations in the NLS and GSOEP. In the IID a clear upward trend was evident; the evidence of an upward trend in the PSID was less clear. The examination of more data is necessary to determine whether this relationship holds up to the test of time.

III. Possible Explanations of the Life Cycle Bias

There are (at least) two reasons to expect estimates of intergenerational earnings persistence to vary with the ages at which fathers' and sons' earnings are measured. First, as discussed in Solon (1989, 1992, 1999) and Mazumder (2001) among others, noise in measured earnings (whether due to mistaken reporting or transitory earnings components) produces an attenuation bias that reduces persistence estimates. A substantial literature (Gottschalk and Moffitt 1994, Buchinsky and Hunt 1999, Gittleman and Joyce 1996, Baker and Solon 2003, Haider 2001) documents a general increase in inequality in both permanent and transitory components in Canada and the US. Growth in transitory earnings variance alone could lead to a larger attenuation bias (and lower persistence estimates) in later time periods; as fathers age, estimates of earnings persistence might diminish. In this explanation the "age-effects" are in actuality "year-effects" since the *year* of observation rather than the age of father determined the

estimated degree of intergenerational earnings persistence.

However, since both permanent and transitory earnings components have increased simultaneously, it is not clear that the attenuation bias in North America has increased over time; if permanent earnings variance has grown more than transitory earnings variance, the attenuation bias has decreased with time and so persistence estimates from recent years contain less downward bias. In Canada, Baker and Solon (2003) find that both permanent and transitory earnings variance increased by similar magnitudes; in the US, Gottschalk and Moffitt (1994) find that transitory variance grew at a rate between 2/3's and equal to that of permanent variance while Haider (2001) finds equal growth. The fact that these studies find that transitory earnings variance grew no faster than permanent earnings variance suggests that attenuation bias has not increases over time and so year effects are not likely explanations for the decrease in estimated earnings persistence as fathers age. Moreover, the fact that earnings persistence seems to have risen as the observation year of sons increased also contradicts the hypothesis that the patterns in the previous section are driven by changes in the degree of attenuation bias.

The theory of human capital accumulation provides a second, unstudied reason for age-dependence in estimates of earnings persistence. Ben-Porath (1967) models life cycle earnings as the outcome of a dynamic investment process. Workers who seek to maximize net lifetime earnings allocate their human capital k between one of two activities: production and learning. In addition, workers purchase investment goods I at price P per unit. Ben-Porath notes that the model is greatly simplified when a Cobb-Douglas form is assumed in the learning technology. In this case, the investment problem faced by the worker is

$$\max_{s(.),i(.)} \int_0^T [R(1-s(t))k(t) - Pi(t)]e^{-rt}dt$$
$$s.t.\ k'(t) = \alpha(s(t)k(t))^{\gamma_1}i(t)^{\gamma_2}$$

(2)

where r is the discount rate, T is the length of the worker's career, R is the rental rate of capital, and s is the fraction of human capital devoted to the learning process. The parameter α represents worker ability which complements learning investments. Given this complementarity, workers with higher ability invest more and so experience greater earnings growth and, ultimately, higher earnings levels. The well-known solution to this investment problem involves large up-front learning investments in both goods and time which gradually diminish over the life cycle as the end of the career draws near. This investment process produces a pattern of increasing earnings variance across the life cycle (or, more accurately, a U-shaped pattern, as Mincer 1974 observed.) Figure 8 illustrates this pattern using a calibrated version of the model drawn from Neal and Rosen (2000).[6] As workers age, the variance in the permanent component of annual earnings rises.

This positive relationship between age and log earnings variance is found in Canada, the US, and Germany as Figures 9, 10, and 11 demonstrate.[7] The data in Figure 9 represent roughly 20,000 fathers drawn from the Canadian IID panel – the same population used in the empirical work of the previous section. As the year of observation is varied from 1978 to 1991 and the fathers age, the variance in log earnings rises by more than 100%. For the American PSID and German GSOEP, similar patterns are shown in Figures 10 and 11. In the PSID and GSOEP, small sample size requires the inclusion of men who are not necessarily fathers. The PSID men are between the ages of 25 and 34 in 1967; the GSOEP men are between the ages of 25 and 43 in

---

[6]The parameter values chosen by Neal and Rosen are $\gamma_1=0.2$, $\gamma_2=0.075$, $r=0.03$, $R/P=4$, $k(18)=1$, $T=65$, and $\alpha=[0.05,0.1,0.15]$.

[7]Clearly, a portion of the observed change may also result from increasing transitory earnings variance.

1983.[8] The only notable qualitative difference in the three graphs is the decrease in log earnings

variance among very young German men. As noted above, Mincer (1974) suggests this variation

should not be alarming.

The effect of such life cycle decision making on estimates of earnings persistence is easy

to see with a simple example.[9] Panel a of Figure 12 simulates the earnings of three individuals.

Solid lines represent the expected paths of log earnings. Transitory shocks cause actual

observations (stars) to deviate from the expected paths. In panel b, the mean trend in the expected

paths and the transitory shocks have been eliminated; this shows the effects of adding age controls

and measurement error correction to a regression like those studied in the previous section. The

example has been constructed consistent with the Ben-Porath model and empirical findings in

Haider and Solon (2003) in that life cycle earnings variance is increasing.

Suppose that fathers and sons from the same family share an expected wage path;

high-earning fathers have high-earning sons while low-earning fathers have low-earning sons.

Once the common trend due to age is subtracted and classical measurement error has been

eliminated, the age-earnings profiles in the population follow the pattern shown in panel b of

Figure 12 with father and son following the same ray. If a lifetime of data were collected for both

---

[8]The maximum age was chosen so that the men were no older than 60 in the last year of
the panel. This was done to mitigate biases that might result as men retired and the composition
of the sample changed. With these age selection rules, the PSID cohort contains roughly 500
men; the GSOEP cohort numbers nearly 1200.

[9]Grawe (2003) develops a formal model following this example. The model is
based on the work of Jenkins (1987), but makes two new contributions. First, Jenkins oddly
assumes that earnings variance decreases with age. Both theory and empirical work suggest the
opposite. More importantly, Jenkins concludes only that changing earnings variance confounds
persistence estimation. This paper (and Grawe 2003) connects known patterns of earnings
variance across the life cycle with clear implications for patterns of earnings persistence
estimates.

fathers and sons, a measurement-error corrected estimate of earnings persistence would be unity. However, suppose only a single annual observation is collected for both father and son. Fix a year in which to observe sons, say at age 30. If sons' log earnings are regressed on fathers' log earnings,

$$y_s = \tilde{\beta} y_f + \varepsilon,$$ (3)

how does the estimate of β change as the year of observation for fathers is varied? For example, consider observing fathers at age 35 versus 55. As the observation point of fathers is moved later in the life cycle, the variance in fathers' earnings grows. A larger variance in fathers' earnings must explain the same variance in sons' earnings; the estimated degree of earnings persistence falls. Similarly, as the observation period for the sons moves later in the life cycle, holding the observation period of fathers constant, the estimated degree of earnings persistence rises. These are exactly the patterns the previous section documents for the United States, Canada, and Germany.

The source of this variation in persistence estimates is a systematic measurement error. While Solon (1989, 1992, 1999) studies the effects of an error that is uncorrelated with fathers' permanent earnings, the analysis of this section suggests that the life cycle accumulation of human capital produces a second measurement error that is strongly correlated with permanent earnings. And, importantly, that correlation varies with age. Once the data are corrected for age, earnings reports from early in the life cycle understate lifetime earnings for individuals with high permanent earnings and overstate lifetime earnings for individuals with low permanent earnings. Later in life, this pattern is reversed. As a result when fathers are observed early in life, persistence estimates contain an upward bias; when fathers are observed later in life, persistence

estimates contain a downward bias. And so we expect to see persistence estimates fall as fathers age.[10] At some (unknown) point in mid-life, the observed annual earning is an unbiased estimate of permanent earnings. Because the measurement error is correlated with the true level of permanent earnings, instrumental variables (IV) will exhibit the same age-dependency as ordinary least squares (OLS); the IV estimates will be higher than OLS estimates, but they will still diminish with as fathers age.

IV. Reconciling a Wide Range of Persistence Estimates

As economists attempt to understand how policy can affect intergenerational earnings persistence, cross-country comparisons are nearly inevitable. Indeed, it is common for authors to compare the persistence estimates generated in their sample with those generated in other countries (see Solon 1992 and Lillard and Kilburn 1995 for two examples). However, whether due to changes in the importance of transitory earnings or life cycle patterns of human capital accumulation, the results above suggest that great care must be taken when comparing estimates of earnings persistence between studies and data sets. In particular, to the extent that samples differ in the average age of fathers (or the year of observation), we should expect samples with older fathers to produce lower persistence estimates. (Since most studies observe sons at roughly the same age – late-20s to mid-30s – the pattern across the sons' life cycles is of limited importance in comparing existing studies. However, despite the small variation in sons ages, the studies summarized in Solon 1999 do show a positive association between average ages of son and estimated persistence.) Moreover, since earnings variance appears to increase over age at an increasing rate we expect the relationship between fathers' ages and persistence estimates to be

---

[10]Clearly, the variation of persistence estimates as sons age can similarly be understood as the result of a systematic measurement error.

concave. What fraction of the variation among studies can be explained by differences in the age of fathers at the point of observation?

Tables 1 and 2 and Figure 13 explore this question, comparing the mean age at which fathers' earnings are measured with the resulting earnings persistence estimates found in all published works that control for attenuation bias following the example of Solon (1992) and include information on fathers' ages at observation. In Table 1, the studies are ordered by the mean age at which fathers' earnings are measured in each study. When a paper includes multiple earnings persistence estimates, the estimate included in Table 1 is the one generated when sample selection rules most closely correspond to the selection rules in Solon (1992) – a) positive annual earnings required in several years which are averaged to control for measurement error and b) include only the oldest son available. The studies by Björklund and Jäntti (1997), Dearden et al. (1997), and Wiegand (1997) are exceptions to the first rule, employing IV methods.

Some studies included in Table 1 do not explicitly report the average age of the fathers in their sample. In these cases, the table reports an approximate age based on other information about the fathers' ages in the paper. In cases in which it is particularly difficult to infer the average age of the father, a question mark follows the range. To the extent that I have erred in approximating the age at which fathers' earnings are observed, I have introduced measurement error into the analysis which presumably serves only to reduce the degree to which the age of fathers at observation can explain differences in intergenerational earnings persistence estimates.

Figure13 plots the reported persistence estimates against average age of fathers; Table 2 summarizes the regression results.[11] As predicted, the relationship between the fathers' ages and

[11]IV estimates are higher by 0.12 on average suggesting either that multi-year averages of fathers' earnings fail to entirely eliminate measurement error or that the instrument is endogenous.

the estimated earnings persistence is strongly negative and concave. A fifteen-year change in the age of the father (from age 40 to age 55) results in a 0.18- to 0.31-point decrease in estimated persistence; this difference is significant in both quadratic ($p=0.013$) and linear models ($p=0.074$). In terms of economic significance, this reduction is enormous. To put it in context, the average earnings persistence estimate in Table 1 is only 0.28. In total, 18 percent of the variance in the estimates can be explained by the error correction methodology (IV as opposed to averaging multiple father earnings observations); of the variance remaining, 35 percent is explained by mean age of father at observation. The method of error correction and the mean age of fathers combine to explain approximately one-half of the existing variation.

In addition to explaining much of the variation between studies, these results substantially alter perceptions of 'outliers' among the studies. For instance, without considering the age of fathers, Couch and Dunn's (1991) persistence estimates of roughly 0.1 for both Germany and the US appear to be extraordinary outliers–especially when the US estimate is compared with other estimates derived from the PSID data used by Couch and Dunn. However, considering the fact that Couch and Dunn observe fathers late in the lifecycle, the results appear in line with other studies.

The correlation between the age of fathers at observation and the year in which fathers were observed ($\rho=0.27$) makes it difficult to differentiate between "age effects" and "year effects"; regressing estimated earnings persistence on both fathers' ages at observation and the year of observation produces negative, statistically insignificant coefficients for both explanatory variables. However, two pieces of information point toward the greater importance of the fathers age. First, more of the variation in earnings persistence is explained by variation in fathers' ages alone than by variation in the year of observation alone (R-square = 0.465 compared with 0.328).

Moreover, it is striking that the age-dependency of earnings persistence estimates across studies (see Figure 13) is nearly the same as that found across years in the American and German data (see Figures 3 and 6). This common trend points away from changes in transitory earnings variance as the explanation for the observed patterns since transitory earnings variance did not increase in all of these countries as it did in the US.

V. An Application to Credit Constraints

A substantial literature has developed to explore the effects of childhood poverty. For example, in *The Consequences of Growing up Poor*, Duncan and Brooks-Gunn (1997) estimate the impact of parental income at different points in the child's life. Using several measures of child success, the authors report large effects when children are young, modest effects when children are young teens, and small to imperceptible effects when children are older teens. Behrman and Taubman similarly (1990) find stronger intergenerational earnings persistence when parents' earnings are measured earlier in the child's life. These authors connect their results to constraints in the credit market that restrict parents' ability to borrow in order to fund their children's educations. (Becker and Tomes 1986 first proposed the connection between earnings persistence and credit market failure.)

The positive relationship between fathers' ages and earnings persistence found in this paper may seem at first to confirm the importance of credit markets. But further examination calls this interpretation into question. First, most of the data in this study cover periods much after the child's education choices are made thus reducing the possible bearing on borrowing constraints that limit education investments. Second, in the case of Germany especially but Canada as well, large government expenditures have been made to reduce the financial expense of these investments. And yet the same patterns across the fathers' life cycles are observed in these

countries and in the US. Finally, in results not printed here, when data are selected to focus on the period just before, during, and after college decisions are made (sons ages 12-18) no notable break is found in the trend of earnings persistence as the son passes the college decision period (roughly ages 15-17).[12] Under the credit constraint explanation, we would expect earnings before the college decision to be distinctly more important than earnings after the decision (even admitting foresight on the part of lenders and parents).

At the very least, the fact that the observed importance of parent earnings in early childhood can easily be explained by a model of human capital accumulation *without* credit constraints underscores the need for more rigorous tests of the credit markets hypothesis. Grawe and Mulligan (2002) survey other attempts to test for constraints.

VI. Conclusion

The development of new panel data sets allows for intergenerational mobility comparisons across countries and groups to identify determinants of earnings persistence. However, the wide variation among published studies hinders such comparative work. This paper finds that a large portion of the variation between earnings persistence estimates can be explained by differences in the age of father at observation. Using data from Canada, Germany, and two data sets from the United States, a strong negative relationship is found between persistence estimates and fathers' ages; a more tenuous positive relationship is found between persistence estimates and sons' ages. Persistence estimates vary by as much as fifty percent as the age at observation shifts by less than fifteen years. Looking across published studies, more than one-third of the variation among studies with similar methodologies can be explained by differences in the age at which fathers are

---

[12]This was done with PSID data. Sample sizes are fairly small with these restrictions limiting the conclusiveness of these results.

observed.

These findings can be explained by two different forms of measurement error. Following Solon (1989, 1992), it may be that increases in transitory earnings variance have produced greater attenuation bias in later years. While plausible, this explanation suffers from two problems. First, while transitory earnings variance has increased in some countries (notably the US), so has permanent earnings variance. Studies report that the increase in permanent earnings variance has been at least as large as that in transitory earnings variance which means the attenuation bias has not increased over time. Second, while the growth in transitory earnings variance has been remarkable in the US, it has not been experienced in other countries like Germany. Yet, the age-dependency of persistence estimates is found in Germany, Canada, and the US.

The second form of measurement error is derived from standard models of human capital accumulation which predict increasing (non-transitory) earnings variance over the life cycle. Given this growth in earnings variance, early observations of earnings are lower than permanent earnings for workers with high permanent earnings and higher than permanent earnings for workers with low permanent earnings (after controlling for age). This measurement error *which is correlated with 'true' permanent earnings* results in lower earnings persistence estimates as fathers age (and higher estimates as sons age). Since the error is correlated with the true value of the independent variable, instrumental variables and other methods for correction for classical measurement error will not eliminate the bias.

Whatever the cause, the strong dependency of persistence estimates on the father's age (or year of observation) alters our understanding of cross-country earnings persistence comparisons. For instance, estimates of persistence in Finland (Österbacka 2001) and Germany (Couch and Dunn 1997) which seemed extremely low appear only slightly lower than expected once the

relatively old age of fathers in these studies is considered. With a fuller understanding of the effects of sampling methodology, we are better prepared to understand the meaning of cross-country differences in persistence estimates and the connects between these differences and policy.

Data Appendix

National Longitudinal Survey

Fathers' wage and salary incomes are recorded in 1966, 1967, 1969, and 1971 for the year prior to the survey. Fathers are restricted to be no older than 55 in 1966 to ensure that a selection bias is not introduced as older fathers retire in later periods. Positive earnings must be reported to be included in the sample.

The sons drawn from the Young Men Cohort are restricted to be no older than 18 in 1966 to avoid oversampling of sons who live at home after high school. Sons' wage and salary incomes from the previous year are reported in 1971, 1973, 1975, 1976, 1978, 1980, and 1981. Given the young age of the respondents, the 1971 and 1973 data are not used. To be included in the sample, the son must report positive earnings. In cases in which more than one son is available from a given household, only the oldest son in the sample is used. Note that this may not be the oldest son in the family since an older son may not have been included in the survey or the sample. The sample sizes range from 270 to 367 depending on the observation years of fathers and sons.

Intergenerational Income Data

The construction of the IID from Canadian tax files is described in detail in Corak and Heisz (1999). The sample studies families with children ages 16-19 in 1982. A one-in-ten sample was taken from the full data set and then, from this sample, the oldest available son for

each family was selected. (Note, the oldest available son may or may not be the oldest son in the family). This resulted in 56,141 father-son pairs. The data was then limited to those fathers born between 1932 and 1942 (inclusive) in order to avoid attrition bias since fathers' labor incomes were recorded from 1978 to 1992. Sons' labor incomes were recorded from 1991 to 1998. The sample includes only observations with positive earnings reports.

Through an examination of the mean and variance of reported incomes, several coding irregularities were found. It appears that a significant number of observations in 1978-1982 were assigned a value of $1 when, in other years, they would have been reported as $0. Similarly, in 1996, a significant number of observations were assigned an earnings of $2. It was not possible to determine why the data included these anomalies. "Positive earnings reports" refer to incomes greater than $1 in 1978-1982 and greater than $2 in 1996.

Panel Study of Income Dynamics

Sons, 9 to 17 years old at the time of the initial 1968 PSID survey, are observed from 1983 to 1992. The exclusion of younger sons ensures that the observations of the sons' incomes are not overly affected by non-representative observations at the beginning of the career. Exclusion of older sons avoids over-representation of sons who live with their parents beyond high school. Since head labor income is used to measure earnings, the son must be the head of household in the observation period in question in order to be included in the sample. Non-positive earnings reports are excluded. In families in which there is more than one son which fits these restrictions, the sample includes only the oldest available son.[13]

"Fathers" in the sample are the male heads of the households in which the sons lived in

---

[13]The study was replicated using the sample of all sons. The results do not change substantially with this alternative sample definition.

1968. They are observed in the years 1967 to 1981. Fathers are eliminated from the sample if their age does not fall between 30 and 46 (inclusive) in 1967. Inclusion of older fathers who will likely retire during the observation period would introduce a sampling bias. Again, fathers must be heads of household in the observation period in question and report positive earnings. The resulting sample sizes range from 199 to 260 depending on the observation years of fathers and sons.

German Socio-Economic Panel

The GSOEP was constructed to mimic the PSID in content and construction. The data for this study are drawn from the West German sample which was collected from 1984 through 2001. Sons, 13 to 17 years old in 1984 are only 30 to 34 in 2001, the final year for which data is available. And so only this one year of earnings data is collected for sons. Exclusion of older sons avoids over-representation of sons who live with their parents beyond high school. Non-positive earnings reports are excluded. Given the very small sample available, the sample was not restricted to only one son per family.

Male heads of household in 1984 form the population of "fathers" in the sample. They are observed from 1984 to 1995. Fathers are eliminated from the sample if their age does not fall between 30 and 46 (inclusive) in 1984. Inclusion of older fathers who will likely retire during the observation period would introduce a sampling bias. Positive earnings in a given year are required to be included in the sample. The sample sizes range from 97 to 127 depending on the observation years of fathers and sons.

References

Altonji, Joseph G. and Dunn, Thomas A. "Family Incomes and Labor Market Outcomes of Relatives." In *Research in Labor Economics*, *Volume 12*, edited by Ronald G. Ehrenberg, 269-310. Greenwich, CT: JAI Press, 1991.

Baker, Michael and Solon, Gary. "Earnings Dynamics and Inequality among Canadian Men, 1976-1992: Evidence from Longitudinal Income Tax Records." *Journal of Labor Economics*, 2003, 21, 289-321.

Ben-Porath, Yoram. "The Production of Human Capital and the Life Cycle of Earnings." *Journal of Political Economy*, August 1967, 75, 352-65.

Björklund, Anders and Markus Jäntti. "Intergenerational Income Mobility in Sweden Compared to the United States." *American Economic Review*, December 1997, 87, 1009-18.

Buchinsky, Moshe and Hunt, Jennifer. "Wage Mobility in the United States." *Review of Economics and Statistics*, August 1999, 81, 351-68.

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

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

Dearden, Lorraine; Machin, Stephen; and Reed, Howard. "Intergenerational Mobility in Britain." *The Economic Journal*, January 1997, 107, 47-66.

Gittleman, Maury and Joyce, Mary. " Earnings Mobility and Long-Run Inequality: An Analysis Using Matched CPS." *Industrial Relations*, April 1996, 33, 181-92.

Gottschalk, Peter and Moffitt, Robert. "The Growth of Earnings Instability in the U.S. Labor Market." *Brookings Papers on Economic Activity*, 1994, 217-72.

Grawe, Nathan D. "Life Cycle Bias in the Estimation of Intergenerational Earnings Persistence." Statistics Canada Analytical Studies Branch Research Paper No. 207, 2003.

Haider, Steven J. "Earnings Instability and Earnings Inequality of Males in the United States: 1967-1991." *Journal of Labor Economics*, October 2001, 19, 799-836.

Haider, Steven J. and Solon, Gary. "Life-Cycle Variation in the Association between Current and Lifetime Earnings." unpublished mimeo, Michigan State University, 2003.

Jenkins, Stephen. "Snapshots Versus Movies: 'Lifecycle Biases' and the Estimation of Intergenerational Earnings Inheritance." *European Economic Review*, 1987, 31, 1149-58.

Lillard, Lee A. And Kilburn, M. Rebecca. "Intergenerational Earnings Links: Sons and Daughters." RAND Working Paper Series, no. 95-17. Santa Monica, CA: RAND Corporation, 1995.

Mazumder, Bhashkar. "Earnings Mobility in the US: A New Look at Intergenerational Inequality." Federal Reserve Bank of Chicago Working Paper, no. 2001-24, 2001.

Mincer, Jacob. *Schooling, Experience, and Earnings*, New York: Columbia University Press (for NBER), 1974.

Mulligan, Casey. *Parental Priorities and Economic Inequality*, Chicago: The University of Chicago Press, 1997.

Neal, Derek and Rosen, Sherwin. "Theories of the Distribution of Earnings," in *Handbook of Income Distribution*, ed. Anthony B. Atkinson and Francis Bourguignon, New York: Elsevier Science, North-Holland, 2000.

Österbacka, Eva. "Family Background and Economic Status in Finland." *Scandinavian Journal*

of Economics, September 2001, 103, 467-84.

Peters, H. Elizabeth. "Patterns of Intergenerational Mobility in Income and Earnings." *Review of Economics and Statistics*, August 1992, 74, 456-66.

Reville, Robert T. "Intertemporal and Life Cycle Variation in Measured Intergenerational Earnings Mobility." RAND Working Paper. Santa Monica, CA: RAND Corporation, 1995.

Shea, John. "Does Parents' Money Matter?" *Journal of Public Economics*, August 2000, 77, 155-84.

Solon, Gary. "Biases in the Estimation of Intergenerational Earnings Correlations." *Review of Economics and Statistics*, February 1989, 71, 172-74.

Solon, Gary. "Intergenerational Income Mobility in the United States." *American Economic Review*, June 1992, 82, 393-408.

Solon, Gary. "Intergenerational Mobility in the Labor Market." In *Handbook of Labor Economics, Volume 3*, edited by Orley Ashenfelter and David Card, 1761-1800. New York: Elsevier, 1999.

Tsai, Shu-Ling. "Sex Differences in the Process of Stratification." Ph.D. dissertation, University of Wisconsin, 1983.

Wiegand, J. "Intergenerational Earnings Mobility in Germany." Mimeo. University College London. 1997.

Zimmerman, David. "Regression Toward Mediocrity in Economic Stature." *American Economic Review*, June 1992, 82, 409-29.

Table 1
Estimates of intergenerational income persistence organized by mean father age.

Author	Mean Age of Father	Mean Year of Father Observation	Estimate	Location
Lillard & Kilburn (1995)	30-40?	1975.5	0.27	Malaysia
Corak & Heisz (1999)	40-45	1980.0	0.23	Canada
Mulligan (1997)	40-45	1969.0	0.33	US
Björklund & Jäntti (1997)*	43	1970.2	0.28	Sweden
Shea (2000)**	44	1969.0	0.36	US
Solon (1992)	44	1969.0	0.41	US
Björklund & Jäntti (1997)*	45	1969.0	0.42	US
Mazumder (2001)+	46	1982.0	0.39	US
Peters (1992)	47	1969.5	0.14	US
Dearden et al. (1997)*	45-50	1974.0	0.58	UK
Tsai (1983)	45-50?	1958.5	0.28	Wisconsin
Österbacka (2001)	48.5	1972.5	0.13	Finland
Couch & Dunn (1997)	51	1986.5	0.11	Germany
Wiegand (1997)*	51	1984.0	0.20	Germany
Altonji and Dunn (1991)	52	1967.3	0.18	US
Couch & Dunn (1997)	53	1986.5	0.13	US

*Note: * Studies using IV estimation.*

**Because Shea does not include the same number of earnings observations for each father, the precise years of observation and the average year of observation are not explicitly stated in the study. However, the data are intended to be very similar to that of Solon.

+Mazumder's estimate using three years of father's data is chosen in order to most closely match the methodology of the other studies in the table. When he uses six years of fathers' earnings data, his estimate is 0.47. His estimates resulting from ten and fifteen years of data are avoided since these estimates were found to be very sensitive to the treatment of top-coded earnings reports.

Table 2
Results from regressing published earnings persistence estimates on age of father at observation.

Effect of Father's Age on Estimates of Earnings Persistence

	Quadratic Model	Linear Model
Father's age (a_f)	0.138 (1.556)	-0.012 (-1.784)
Father's age^2 (a_f^2)	-0.002 (-1.692)	NA
IV dummy	0.109 (1.725)	0.130 (1.949)
$E[\beta|a_f=40]-E[\beta|a_f=55]$	0.305 (2.549)	0.175 (1.784)
R-square	0.465	0.338

Note: t-values in parenthesis.

Figure 1
**Pattern of earnings persistence across year of father observation in the NLS.**
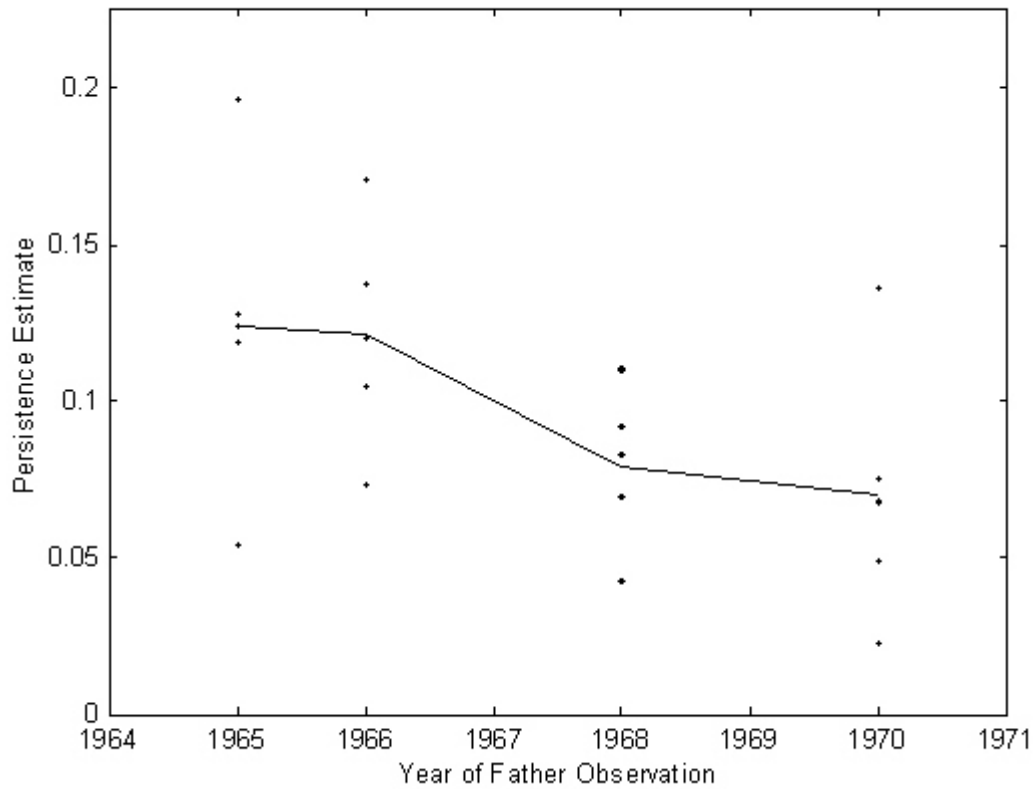
Note: The NLS surveys of 1966, 1967, 1969, and 1971 requested data on fathers' earnings in the previous year. Surveys in 1975, 1976, 1978, 1980, and 1981 requested data on sons' earnings in the previous year. The four observations of fathers and five observations of sons can be paired in 20 different combinations to create 20 different estimates of intergenerational earnings persistence. Each of these 20 different earnings persistence estimates is represented by a point in the figure. For each year of father observation, the five resulting persistence estimates are averaged; the solid line shows how this average varies across the year of father observation.

Figure 2
**Pattern of earnings persistence across year of son observation in the NLS.**
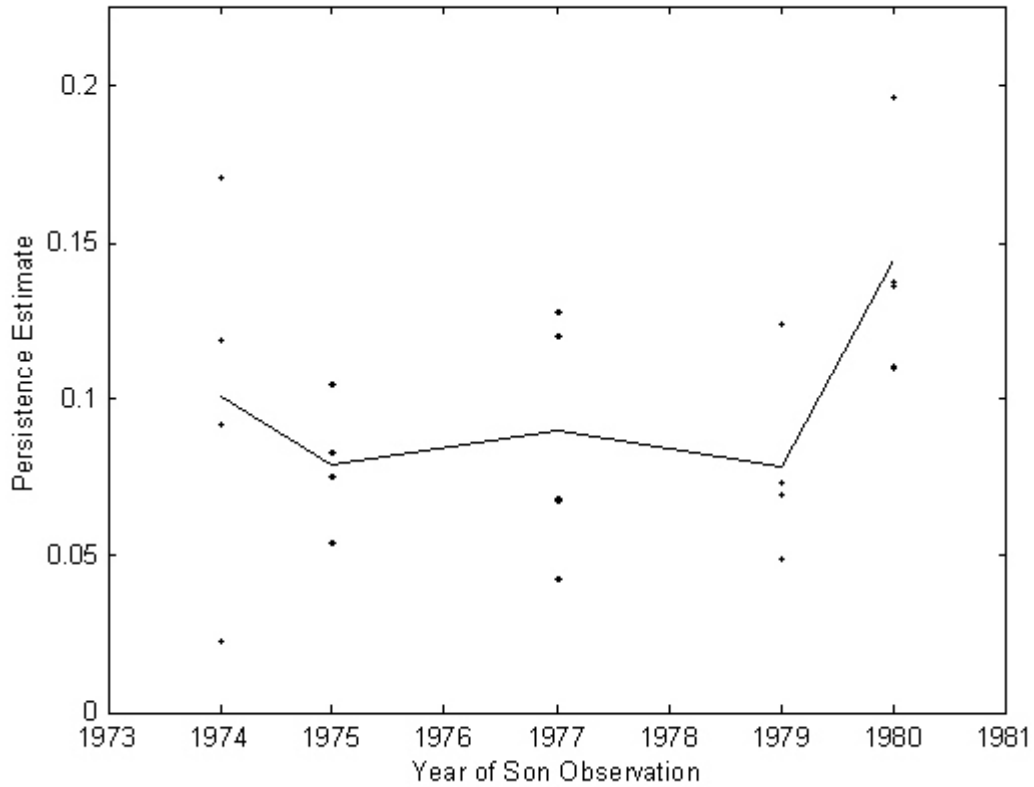
Note: The NLS surveys of 1966, 1967, 1969, and 1971 requested data on fathers' earnings in the previous year. Surveys in 1975, 1976, 1978, 1980, and 1981 requested data on sons' earnings in the previous year. The four observations of fathers and five observations of sons can be paired in 20 different combinations to create 20 different estimates of intergenerational earnings persistence. Each of these 20 different earnings persistence estimates is represented by a point in the figure. For each year of son observation, the four resulting persistence estimates are averaged; the solid line shows how this average varies across the year of son observation.

Figure 3
Pattern of earnings persistence across year of father observation in the GSOEP.


Note: The GSOEP surveys from 1984-1998 requested data on fathers' earnings in the previous year. Sons' 1999 earnings were observed in 2000. Surveys in 1975, 1976, 1978, 1980, and 1981 requested data on sons' earnings in the previous year. The 15 observations of fathers are paired with the observation of sons' 1999 earnings to create 15 different estimates of intergenerational earnings persistence. The solid line shows how these 15 persistence estimates varies across the year of father observation.

Figure 4
**Pattern of earnings persistence across year of father observation in the IID.**
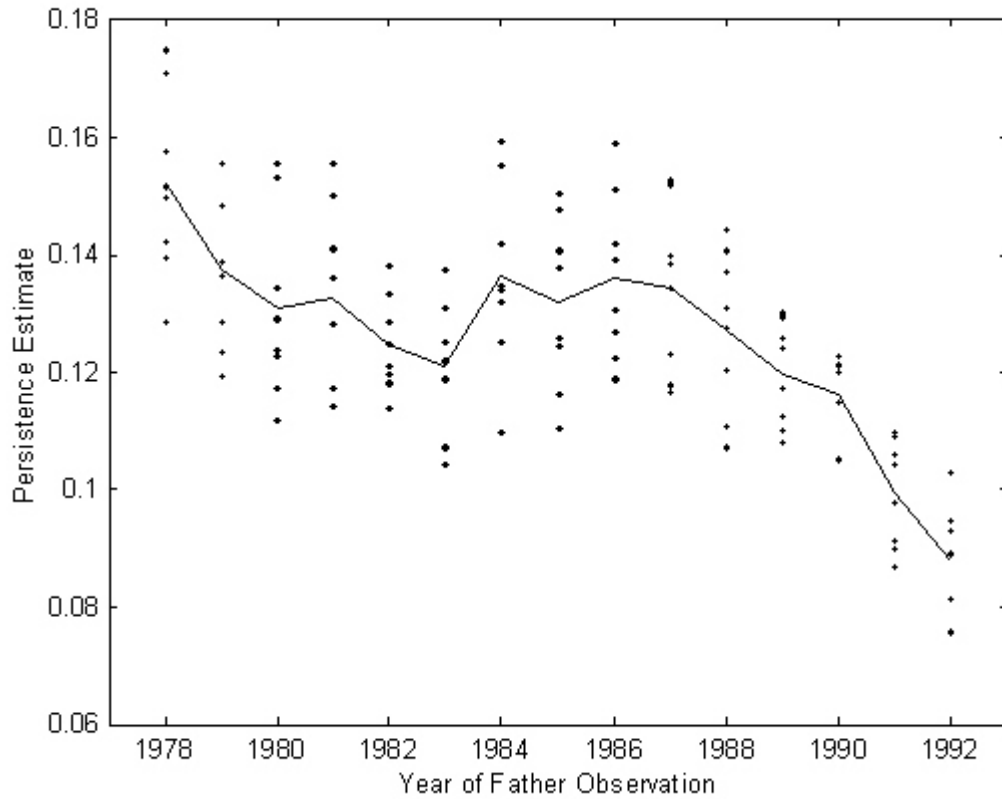
Note: The IID includes data on fathers' earnings from 1978 to 1992 . Sons' earnings are observed from 1991 to 1998. The 15 observations of fathers and eight observations of sons can be paired in 120 different combinations to create 120 different estimates of intergenerational earnings persistence. Each of these 120 different earnings persistence estimates is represented by a point in the figure. For each year of father observation, the eight resulting persistence estimates are averaged; the solid line shows how this average varies across the year of father observation.

Figure 5
**Pattern of earnings persistence across year of son observation in the IID.**
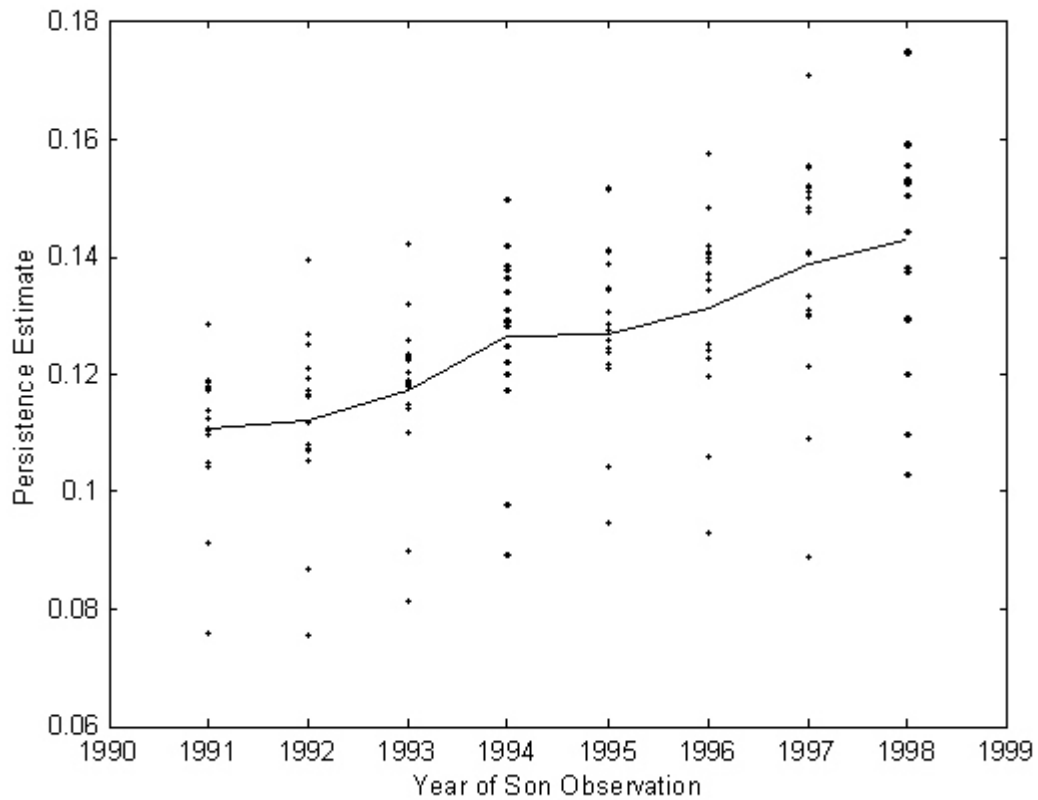
Note: The IID includes data on fathers' earnings from 1978 to 1992 . Sons' earnings are observed from 1991 to 1998. The 15 observations of fathers and eight observations of sons can be paired in 120 different combinations to create 120 different estimates of intergenerational earnings persistence. Each of these 120 different earnings persistence estimates is represented by a point in the figure. For each year of son observation, the 15 resulting persistence estimates are averaged; the solid line shows how this average varies across the year of son observation.

Figure 6
Pattern of earnings persistence across year of father observation in the PSID.
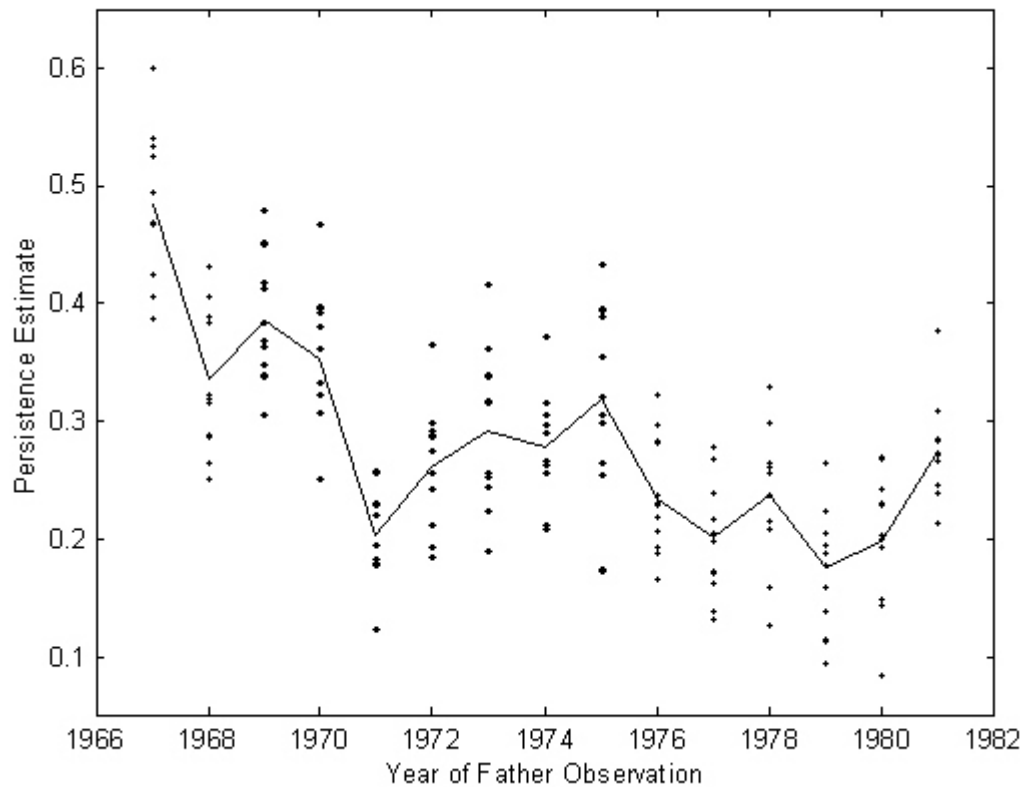
Note: The PSID from 1968 to 1982 includes data on fathers' earnings in the previous year. Sons' earnings from the previous year are observed from 1984 to 1993. The 15 observations of fathers and ten observations of sons can be paired in 150 different combinations to create 150 different estimates of intergenerational earnings persistence. Each of these 150 different earnings persistence estimates is represented by a point in the figure. For each year of father observation, the ten resulting persistence estimates are averaged; the solid line shows how this average varies across the year of father observation.

Figure 7
Pattern of earnings persistence across year of son observation in the PSID.
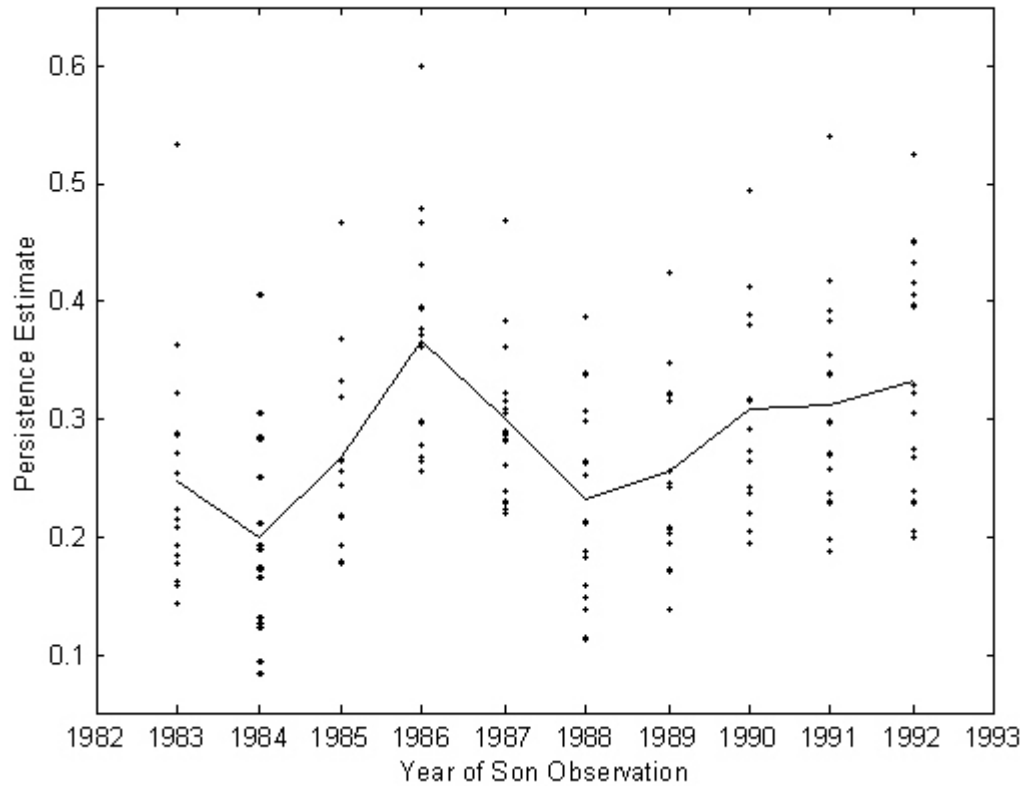
Note: The PSID from 1968 to 1982 includes data on fathers' earnings in the previous year. Sons' earnings from the previous year are observed from 1984 to 1993. The 15 observations of fathers and ten observations of sons can be paired in 150 different combinations to create 150 different estimates of intergenerational earnings persistence. Each of these 150 different earnings persistence estimates is represented by a point in the figure. For each year of son observation, the 15 resulting persistence estimates are averaged; the solid line shows how this average varies across the year of son observation.

Figure 8
Life cycle pattern in earnings in a Ben-Porath model.
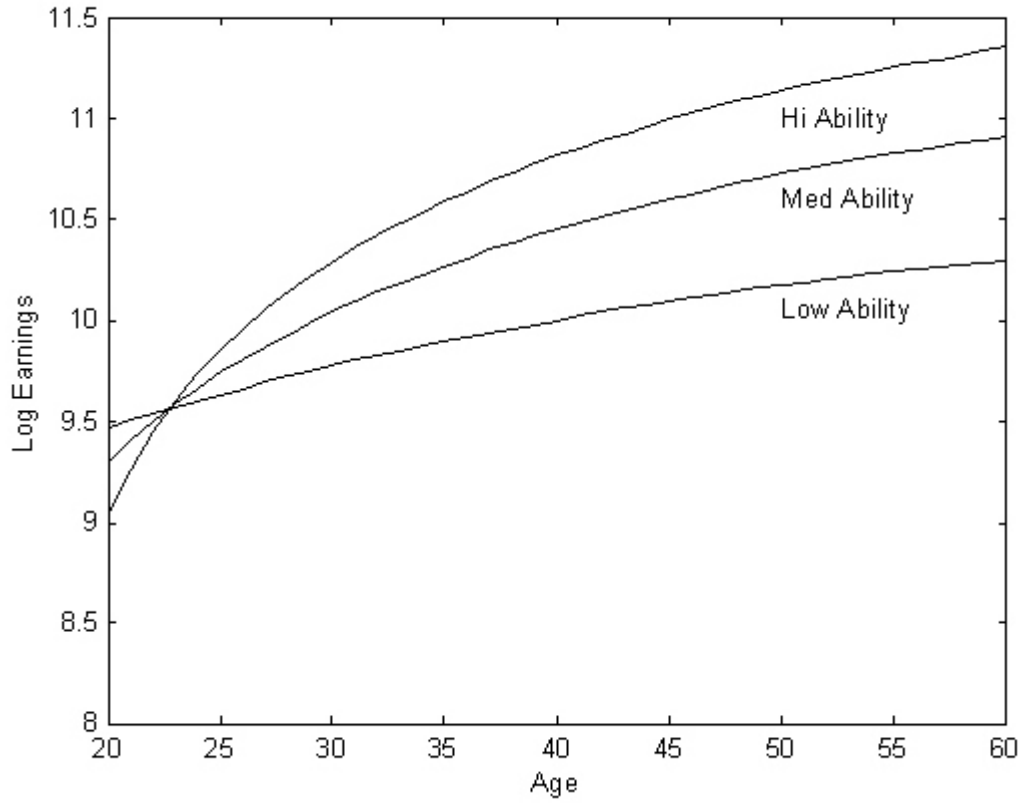
Note: Simulations follow Neal and Rosen (2000) with parameter values $\gamma_1=0.2$, $\gamma_2=0.075$, $r=0.03$, $R/P=4$, $k(18)=1$, $T=65$, and $\alpha=[0.05,0.1,0.15]$.

Figure 9
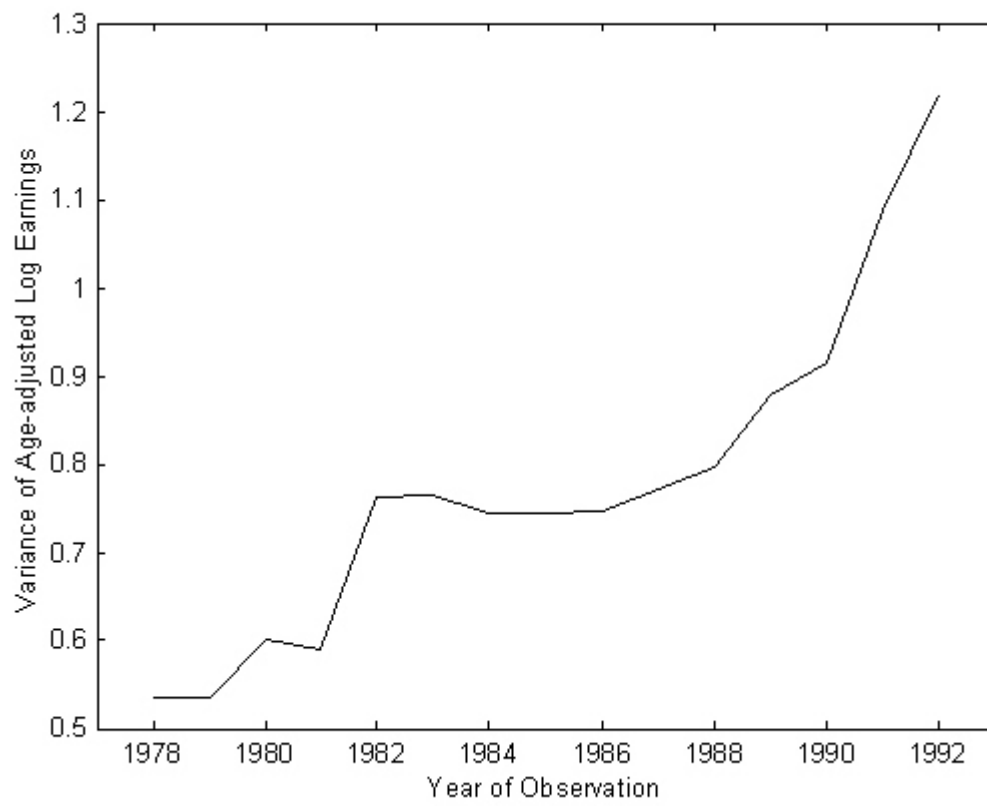**Variance of annual earnings over the life cycle of Canadian IID fathers.**

Figure 10
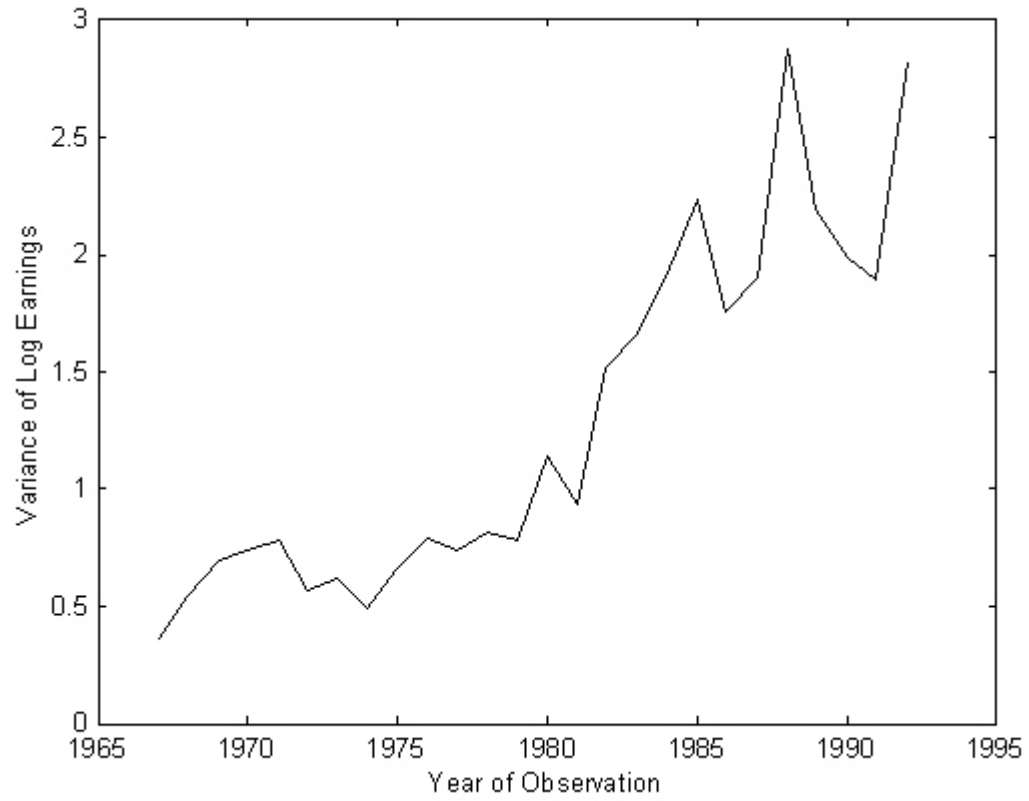**Variance of log annual earnings over the life cycle of American PSID males ages 25-34 in 1967.**

Figure 11
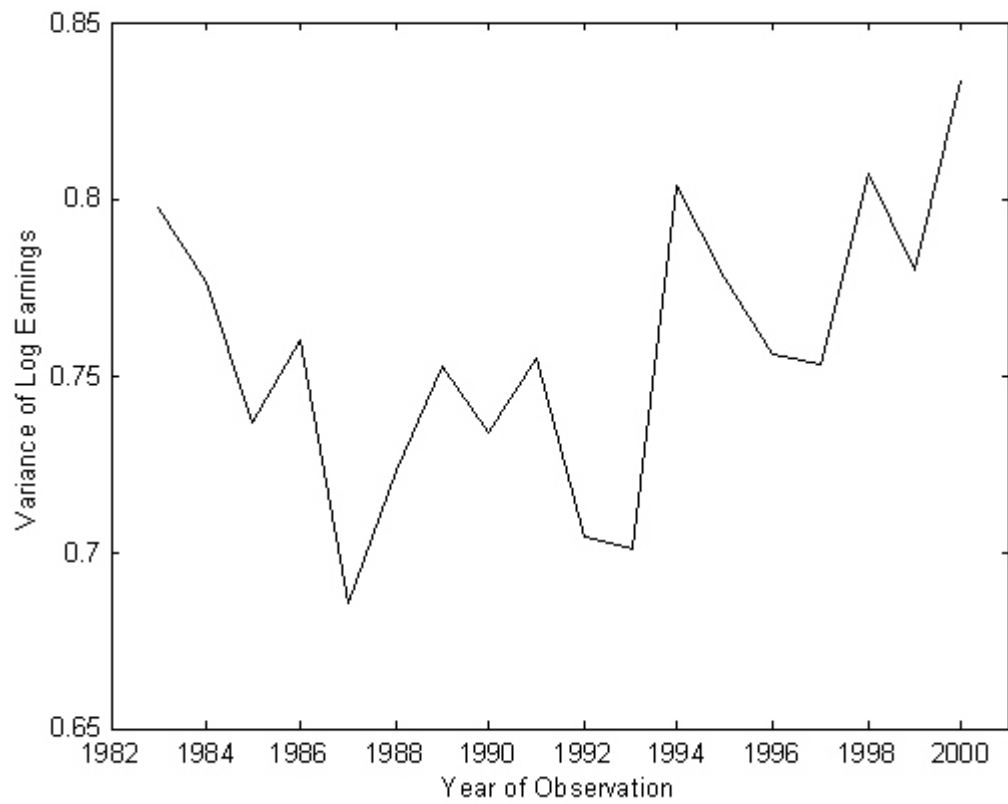**Variance of log annual earnings over the life cycle of American GSOEP males ages 25-43 in 1983.**
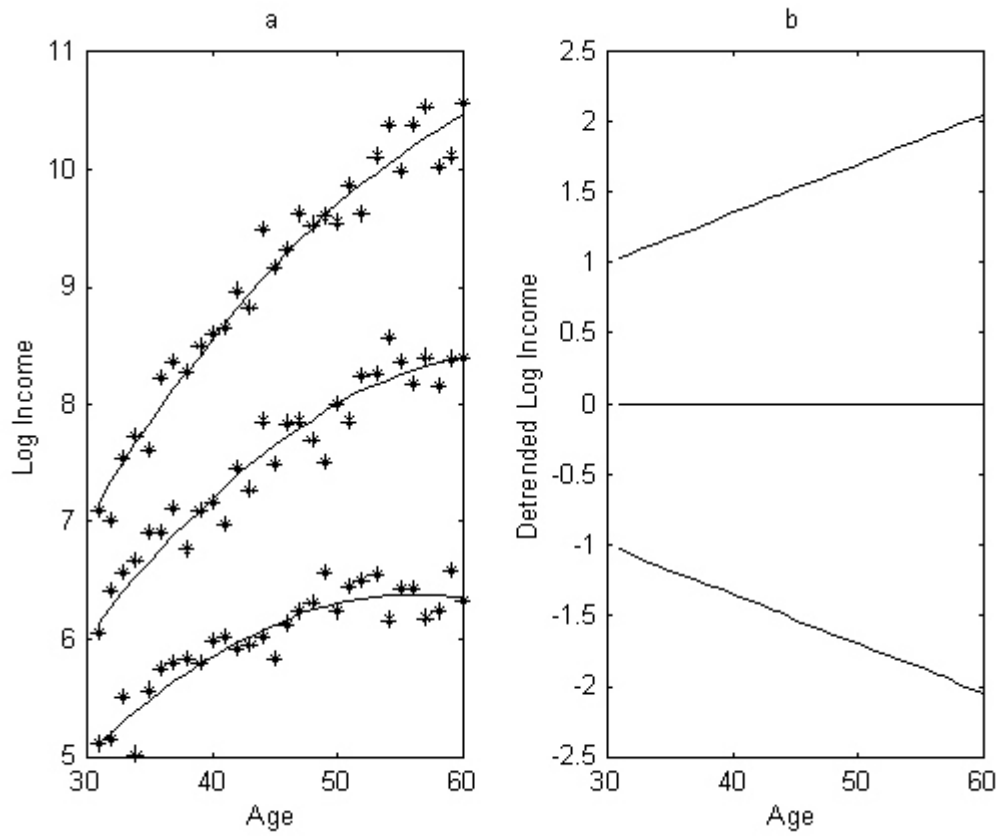
Figure 12
Simulated age-income profiles before and after detrending and measurement error correction.

Note: Solid curves represent expected earnings paths; stars represent actual observations. a, Raw age-income profiles. b, Detrended, error-purged age-income profiles.

Figure 13
Earnings persistence estimates from published studies plotted against age of fathers at observation.
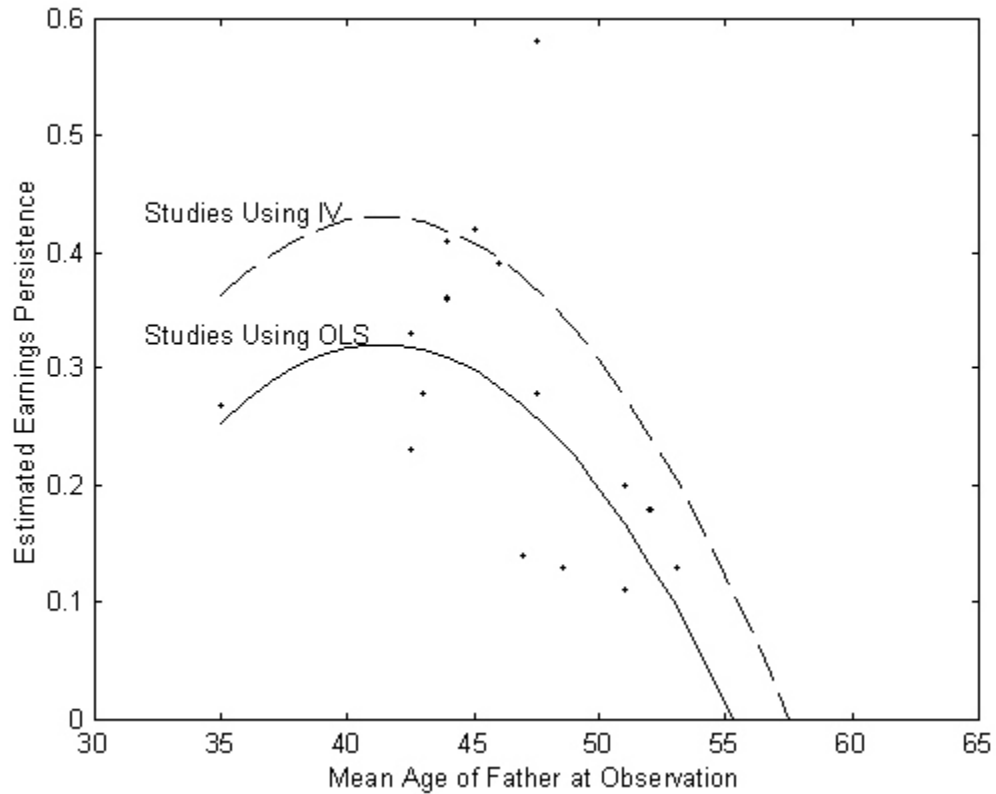
Note: Lines plot the expected persistence estimate given fathers' mean age, The solid line shows the expected earnings persistence for a study using ordinary least squares to estimate the relationship between log earnings of the son and log earnings of the father where the father's earnings are measured as the average earnings over several years to account for classical measurement error. The dashed line shows the expected earnings persistence for a study using instrumental variables to estimate the relationship between log earnings of the son and log earnings of the father.