

# The Intergenerational Persistence of Lifetime Earnings\*

Lutz Hendricks  
Arizona State University  
Department of Economics  
February 5, 2002

## Abstract

This paper proposes a new method for estimating the intergenerational persistence of lifetime earnings from data that contain only short sections of individual earnings histories. The approach infers lifetime earnings persistence from the persistence of short earnings averages together with information about the stochastic process governing individual earnings. I find that lifetime earnings are substantially more persistent than estimates of average earnings persistence suggest. The coefficient in a regression of children's lifetime earnings on fathers' lifetime earnings is approximately 0.54. Proxying for lifetime earnings using five year averages leads to a downward bias in estimated intergenerational persistence of one-third. The bias is much stronger, if observations with zero earnings are not excluded from the sample. These findings are robust against alternative assumptions about the data generating process for earnings.

## 1 Introduction

This paper studies the measurement of intergenerational earnings persistence. Several indicators of economic status, such as income and wealth, are transmitted from parents to their children. This raises concerns about equality of opportunities and motivates numerous public policies designed to foster the economic outcomes of the poor. Intergenerational persistence of economic status may also be an important source of inequality.

A common measure of intergenerational persistence is the coefficient  $\rho_c$  in a regression of the form

$$\ln E_c(i) = \theta + \rho_c \ln E_c(p(i)) + v(i) \quad (1)$$

where  $E_c(i)$  is an indicator of person  $i$ 's economic status (such as earnings or wealth),  $v(i)$  is a random shock, and  $p(i)$  denotes  $i$ 's parent. If  $\rho_c = 1$ , then differences in status are perfectly transmitted from parents to children (complete persistence), whereas if  $\rho_c = 0$ , then child outcomes are uncorrelated with parental outcomes (complete mobility). In spite of numerous efforts to estimate  $\rho_c$  for lifetime earnings, no consensus has emerged as to whether intergenerational persistence is weak or strong. Recent estimates of  $\rho_c$  range from almost complete mobility ( $\rho_c = 0.05$  in Couch and Lillard 1998) to strong persistence ( $\rho_c = 0.5$  in Solon 1992). Since these findings are obtained from the same data, the very different persistence estimates must be due to differences in estimation methods.

The central difficulty in estimating  $\rho_c$  is that (1) should be estimated for *lifetime* earnings. In applications, lifetime earnings are commonly defined as a discounted present value of earnings over

---

\*I am grateful to Seung Ahn for his advice on econometric issues. Financial support from the College of Business at Arizona State University is gratefully acknowledged.

a long age range (e.g., Fullerton and Rogers 1993; Gokhale et al. 2001). However, since commonly used datasets contain only short sections of individuals' earnings histories, lifetime earnings are not observable. This raises the question how lifetime earnings persistence can be inferred from available data.

Previous research addressed this problem by decomposing log earnings at age  $a$  into a permanent component ( $h$ ) and a transitory component ( $z$ ):<sup>1</sup>

$$y(i, a) = h(i) + z(i, a)$$

The permanent component is transmitted from parents to children according to

$$h(i) = \beta h(p(i)) + \eta(i) \tag{2}$$

where  $\eta(i)$  is a random shock. Equation (2) is then estimated using a measure of  $h$  and the resulting estimate of  $\beta$  is interpreted as lifetime earnings persistence.<sup>2</sup> Three common measures of  $h$  are used: (i) average earnings over up to 5 years, (ii) instrumental variable estimates of  $h$ , and (iii) the individual fixed effect in a panel regression.

The point of this paper is that the resulting estimates of  $\beta$  do not consistently estimate lifetime earnings persistence. For plausible earnings processes, the intergenerational persistence of  $h$  can be very different from that of lifetime earnings. Therefore, a new estimation method is proposed that avoids equating permanent earnings ( $h$ ) with lifetime earnings ( $E_c$ ).

What motivates the use of  $h$  as a proxy for lifetime earnings is intuition that the transitory shocks average out over a large number of years. An average of  $y(i, a)$  is then (approximately) equal to  $h(i)$  (plus a constant), and replacing  $E_c(i)$  with  $h(i)$  in (1) does not introduce a bias ( $\beta = \rho_c$ ). This intuition is correct, if the  $z$  shocks are not too large and not too persistent. However, empirical estimates of earnings processes indicate that  $z$  is highly persistent and accounts for a large part of the variance of log earnings around middle age (details below). Then the  $z$ 's do not average out and the persistence of lifetime earnings can be very different from that of  $h$ .

To address this problem, this paper proposes a new estimation approach which avoids equating lifetime earnings with  $h$ . Instead, it explicitly models the intergenerational persistence of lifetime earnings as a function of the parameters governing the data generating process (DGP) for earnings. Specifically, the approach involves three steps.

Step 1 estimates parameters of the earnings DGP not related to intergenerational persistence. Many such estimates are reported in the literature.

Step 2 estimates the intergenerational persistence parameters of a given earnings DGP using a simple method of moments approach. Essentially, these parameters are chosen to match intergenerational persistence of short earnings averages that can be observed in available data.

Step 3 draws a large number of parent-child pairs and computes the intergenerational persistence of lifetime earnings.

This approach has several benefits. (i) The information from various estimation methods and samples can be combined to arrive at the final estimate of  $\rho_c$ . At the same time the approach clarifies the sources of differences between alternative estimation methods proposed in the literature. (ii) A major concern in the literature is whether different earnings averages or instrumental variable measures provide good proxies for  $h$ . No such proxies are needed here because the relationship between observed

---

<sup>1</sup>Here I adopt the language of the intergenerational persistence literature, which differs from that of the econometrics literature. Note, in particular, that  $z$  could have a unit root.

<sup>2</sup>This method is stated most clearly in Zimmerman (1992), but underlies many other studies. One exception is Knowles (1999) who explicitly estimates the persistence of lifetime earnings, albeit using a very different approach from the one proposed here.

objects (e.g., earnings averages) and lifetime earnings is modeled explicitly. *(iii)* The estimate of  $\rho_c$  can be computed to precisely match any definition of lifetime earnings. This is an important benefit because studies differ in the details of how the present value of lifetime earnings is defined (e.g., based on full-time or realized earnings over different age ranges). *(iv)* The approach can be extended to estimate the intergenerational persistence of other variables for which only short sections of individual histories are observable. Examples include consumption, income, and wages.

An apparent drawback of the proposed approach is that the conclusions depend on the assumed earnings DGP. However, other methods of estimating intergenerational persistence, of course, share this drawback; it is merely not as apparent if the DGP is not explicitly stated. Fortunately, the findings turn out to be very similar across alternative DGPs within a class of commonly estimated processes. Moreover, the assumed DGP may be validated by requiring that it replicates the persistence estimates obtained according to various methods proposed in the literature.

Based on studies that use intergenerational persistence estimates to parameterize general equilibrium models, I consider two definitions of lifetime earnings: the discounted present value of realized lifetime earnings ( $E_P$ ) and the discounted present value of full-time or potential earnings ( $E_F$ ). Both lifetime earnings concepts measure the discounted present value of earnings over a large number of years. They differ in the age ranges covered and in the treatment of years during which a person works part time or not at all.  $E_P$  covers the entire lifetime and retains observations with zero or part-time earnings. This measure is used, for example, in studies of saving or intergenerational transfer behavior, such as Gokhale et al. (2001).  $E_F$  measures a person's earnings potential over a fixed range of years. Observations with zero or part time earnings are interpolated.  $E_F$  is used to measure the distribution of lifetime labor endowments (e.g., Fullerton and Rogers 1993).

The main findings are as follows. Typical estimates of the intergenerational persistence of 5 year average earnings are near 0.37. Earnings processes that replicate such estimates imply values of lifetime earnings persistence near 0.54 for  $E_P$  and  $E_F$ . In other words, using average earnings to proxy for lifetime earnings entails a downward bias of around one-third. The intuition is that a 5 year window is not long enough to eliminate attenuation bias, given the high persistence of earnings shocks and given that parents are typically observed at ages past 40.

Consistent with the findings of Couch and Lillard (1998), estimated persistence from average earnings falls to roughly 0.1 if observations with zero earnings are not deleted. If the objective is the estimation of lifetime earnings persistence, not deleting or interpolating atypical observations with zero or very low earnings greatly exacerbates the bias of the estimator. Matching estimates of average earnings persistence near 0.37 requires that the permanent component of earnings ( $h$ ) is highly persistent. All of these findings are robust against variations in the earnings DGP within a class of models commonly used in the macroeconomics literature.

Some studies have proposed instrumental variable estimators as a way of separating permanent from transitory earnings components. I argue that such estimators are potentially useful for estimating the persistence of  $h$ , but contain little additional information about the persistence of lifetime earnings as measured by  $E_P$  or  $E_F$ .

The remainder of the paper is organized as follows. Section 2 reviews existing empirical estimates of intergenerational earnings mobility and their underlying assumptions. Section 3 develops an empirical model of earnings dynamics. The estimation approach is described in section 4. The findings are presented in section 5, and section 6 concludes.

## 2 A Review of Empirical Persistence Estimates

This section selectively reviews previous empirical estimates of intergenerational earnings persistence. No attempt is made to offer a comprehensive survey (see Mulligan 1997 and Couch and Lillard 1998

for more comprehensive reviews). Instead, I select studies that highlight how alternative estimation methods affect the conclusions about intergenerational persistence. Table 1 summarizes these studies.

It is helpful to establish notation first. I denote an  $m$  year average of log earnings by  $E_m$  and its intergenerational persistence by  $\rho_m$ . Instrumental variable estimates of lifetime earnings are denoted  $E_{IV}$  with a persistence coefficient of  $\rho_{IV}$ . Proxying for lifetime earnings based on a fixed effect panel regression (Knowles 1999) yields  $E_K$  with persistence  $\rho_K$ .

Early studies, such as Behrman and Taubman (1985), estimated (1) based on earnings in a single year. These studies resulted in estimates of  $\rho_1$  on the order of 0.25. The innovation of Solon (1992) and Zimmerman (1992) was to recognize that OLS estimates suffer from attenuation bias. This bias stems in part from nonrepresentative samples and in part from measurement error and transitory shocks. Using more representative samples and proxying lifetime earnings with average earnings over up to 5 years, both authors arrive at persistence estimates on the order of 0.35 to 0.4. Couch and Lillard (1998) argue that these higher estimates are biased upward because of restrictive sample selection criteria, especially dropping observations with zero or low earnings. If all individuals are included in the sample, much lower estimates of  $\rho_m$ , typically below 0.1, are obtained. A similar estimate is reported by Peters (1992) who retains all parents and sons with at least one strictly positive earnings observation ( $\rho_5 = 0.14$ ).

Two alternative estimation techniques avoid the use of average earnings as a proxy for lifetime earnings. The first approach uses instrumental variables, such as education or occupation, to estimate parental permanent earnings (in the sense of  $h$ ). The resulting persistence estimates tend to be larger than those obtained from earnings averages, but they are somewhat sensitive to the choice of instruments, ranging from 0.36 (Zimmerman 1992) to 0.53 (Solon 1992).

The second alternative approach estimates the present value of lifetime earnings using a fixed effect panel regression. Knowles (1999) defines a *household's* lifetime earnings as the present discounted value of head's and spouse's earnings between the ages of 30 and 80. In years where a household's earnings are zero or not observed, they are interpolated based on an estimated age-earnings profile. Hence, a household's lifetime earnings are approximately proportional to its fixed effect in the panel regression. The resulting estimate of lifetime household earnings persistence is 0.34. Hendricks (2001) applies a similar procedure to individual earnings data and arrives at an estimates of  $\rho_F$  near 0.3.

This selective review shows that previous studies differ in their proxies for lifetime earnings (typically  $E_1$  or  $E_5$ ), in sample selection criteria, and in the treatment of observations with zero earnings. In addition, the estimation approaches differ in their data sources (NLS or PSID) and in the ages at which parents and children are observed (see table 1). In what follows, I shall take the estimates of  $\rho_m$  for each estimation method as representative of other estimates in the literature. In particular, I shall assume that  $\rho_5 = 0.37$  for 5-year average earnings where parents and children are observed at ages around 47 and 33, respectively, and observations with zero earnings are deleted. These figures capture the means of Solon's and Zimmerman's samples and estimates. Similar estimates are obtained based on proxies for lifetime earnings, which are in turn estimated from roughly 5-year sections of individual earnings histories (e.g., Knowles 1999). Finally, not deleting observations with zero earnings reduces  $\rho_5$  to roughly 0.1, based on Couch and Lillard (1998). While individual studies differ slightly from these estimates, the rough magnitudes appear to characterize a large fraction of the existing evidence.

[INSERT TABLE 1 HERE]

### 3 An Empirical Model of Labor Earnings

This section develops a statistical model of labor earnings that nests several models estimated in the literature. Individuals are indexed by  $i$  and participate in the labor market between the ages of  $a_1$  and  $a_R$ . Denote the logarithm of (latent) earnings of person  $i$  at age  $a$  by  $y(i, a)$ . This evolves according to

the process

$$y(i, a) = g(a) + h(i) + z(i, a) + \varepsilon(i, a) \tag{3}$$

where

$$\begin{aligned} h(i) &= \beta h(p(i)) + \eta(i) \\ z(i, a + 1) &= \alpha z(i, a) + \zeta(i, a + 1) \\ z(i, a_1) &= \varphi z(p(i), a_B) + \omega(i) \\ \varepsilon(i, a) &\sim N(0, \sigma_\varepsilon^2) \\ \eta(i) &\sim N(0, \sigma_\eta^2) \\ \omega(i) &\sim N(0, \sigma_\omega^2) \end{aligned}$$

Log earnings consist of the following components:

1. A deterministic age profile,  $g(a)$ .
2. A permanent component  $h(i)$  with intergenerational persistence  $\beta$ .
3. A persistent component,  $z(i, a)$ , which evolves as an AR(1) with persistence  $\alpha$ .
4. A transitory shock,  $\varepsilon(i, a)$ , which is drawn *iid* from a Normal distribution.

The disturbances  $(\varepsilon, \eta, \zeta, \omega)$  are mutually independent. Each person has one parent, indexed by  $p(i)$ . Upon entering the model, a person draws the endowment  $z(i, a_1)$  which depends on the parent's  $z$  at age  $a_B$ . Realized earnings are given by

$$Y(i, a) = \exp(y(i, a)) I(e(i, a) = 1)$$

where  $I$  is an indicator function and  $e(i, a)$  denotes an employment shock which equals one with probability  $\pi(a)$  and zero otherwise. When unemployed, earnings equal zero; when employed, log earnings are given by  $y(i, a)$ . Earnings processes that fit into the framework of equation (3) have been estimated numerous times. Examples are provided in table 2.

[INSERT TABLE 2 HERE]

The model abstracts from a number of features that might be of interest. Education is persistent across generations and should affect the mean profile  $g(a)$  and the date of entry into the labor market,  $a_1$ . Similarly, the date of exit from the labor force,  $a_R$ , varies across individuals. Earnings could be modeled as the product of wages and hours worked. This would permit to model the effect of sample selection rules based on hours worked. In the data, the probability of being unemployed probably depends on latent earnings as well as on past unemployment shocks. The joint effect of mother's and father's earnings on child earnings could be considered in a two-parent model, although this is not commonly done in the empirical literature. I do not consider these features, mainly in order to remain consistent with previous research. However, the estimation approach proposed here could accommodate them.

## 4 Estimation

Before describing the estimation approach, it is useful to discuss a number of model features that are relevant for estimation. In the data,  $h(i)$  may be thought of as individual characteristics that are fixed early in life, such as aptitudes, intelligence, or education. In a more complete model,  $\beta$  and  $h(i)$  would

be vectors so that different characteristics may differ in their intergenerational persistence. Also note that earnings are not directly persistent. Instead, earnings are a function of individual characteristics, which are transmitted from parents to children. Hence, (1) is not a structural equation but a reduced form, and  $\rho_c$  is not a structural parameter but an unknown function of  $\beta$  and  $\varphi$ .

This raises the issue of exactly how  $\rho_c$  is to be interpreted. One interpretation is that  $\rho_c$  answers the question: "By how much does an exogenous one percent increase in parental lifetime earnings increase mean child lifetime earnings?" This notion of  $\rho_c$  is not well-defined. To what extent parental earnings changes are transmitted from parents to children depends on what causes these changes. A unit increase in  $h(p(i))$  raises child earnings by  $\beta$ , whereas a unit increase in  $\varepsilon(p(i), a)$  has no effect on child earnings. The fact that multiple characteristics are likely intergenerationally persistent ( $\beta$  is a vector) amplifies this ambiguity. For example, a positive earnings shock received by the parent late in life due to pure luck likely has a smaller effect on child earnings than a parental earnings increase due to additional education. This observation has important implications for instrumental variable methods as discussed further in section 5.2.2.

My interpretation of  $\rho_c$  is motivated by the way estimates of intergenerational persistence are used in the literature. One common use is the parameterization of earnings processes in computable general equilibrium models (e.g., Fullerton and Rogers 1993; Castaneda et al. 2000; Hendricks 2001). In these applications,  $\rho_c$  may be thought of as answering the question: "If a *randomly chosen* parent earns one percent more than the mean, then the child earns, on average,  $\rho_c$  percent more than the mean." This interpretation is also suitable as a descriptive statistic for intergenerational persistence. As I shall argue below, a number of estimation methods employed in earlier research answer a different question. Of course, this does not invalidate the answers; however, it makes the estimates unsuitable for calibrating earnings processes in general equilibrium models.

The remainder of this section describes the estimation approach. It proceeds in three steps, which are described in detail in the following three subsections.

#### 4.1 Parameterizing the Earnings Process

The first estimation step determines the parameters of the earnings process not related to intergenerational persistence. These are summarized in table 3. Workers enter the labor market at age  $a_1 = 23$  and exit at age  $a_R = 65$ . The age-earnings profile,  $g(a)$ , is set to median male earnings in 1985 taken from the Social Security Bulletin (1998, table 4B). The intercept,  $g(a_1)$ , only matters for replicating some sample selection rules used in the literature. It is set to match mean log earnings of \$28,000 at age 42 for parents and of \$41,000 at age 29 for sons. These are the averages of ages and earnings in Couch and Lillard's (1998) PSID sample. If the earnings distribution is log-normal, then mean log earnings equal median earnings.

The probability of zero earnings,  $\pi(a)$ , is set to equal the fraction of male individuals of age  $a$  in the PSID who report zero labor income. These range from 2% at age 25 to 29% at age 65. The low fractions at early age may be due to the fact that persons with zero labor income do not head their own households and are therefore not captured in the PSID. The benefit of using  $\pi(a)$  derived from the PSID is consistency with the data used by Couch and Lillard (1998).

[INSERT TABLE 3 HERE]

The remaining parameters characterize the evolution of earnings over a person's lifetime and are taken from previous empirical studies. The baseline case is based on Storesletten et al. (1998, hereafter STY). The parameters of the DGP are estimated based on a stochastic life-cycle model using the Generalized Method of Moments. Essentially, the parameters match the variance of log earnings at various ages in a sequence of cross-sectional earnings distributions taken from the PSID. The key identifying assumption is that new agents start out with the same persistent shock:  $z(i, a_1) = 0$ . The robustness of the findings to alternative specifications of the earnings process are considered in section

5.1.

## 4.2 Estimating Intergenerational Persistence Parameters

The intergenerational persistence parameters ( $\beta$  and  $\varphi$ ) are estimated using a simple method of moments approach. Most earnings DGPs estimated in the literature permit only one source of intergenerational persistence. In STY's case,  $z(a_1, i) = 0$ , so that  $\varphi$  becomes irrelevant. The other DGPs set  $h = 0$ , so that  $\beta$  becomes irrelevant.<sup>3</sup> Then a single parameter, either  $\beta$  or  $\varphi$ , determines intergenerational persistence. It is chosen such that regressing average child earnings at ages 30 to 33 on average parental earnings at ages 45 to 48 yield a slope coefficient of  $\rho_5 = 0.37$ . This approximately replicates the findings of Solon (1992) and Zimmerman (1992). The age at which  $z$  is transmitted from parent to child is set to  $\alpha_B = 35$ , which corresponds to the middle of childhood for a typical child.

The sensitivity analysis considers cases where both  $h$  and  $z(a_1)$  are transmitted from parent to child. It would be possible to identify  $\beta$  and  $\varphi$  by matching two or more intergenerational persistence observations. However, the sensitivity analysis shows that alternative combinations of  $\beta$  and  $\varphi$  that yield the same value of  $\rho_5$  are nearly observationally equivalent. As a result, identification would be weak. I therefore explore all possible combinations of  $\beta$  and  $\varphi$  that are consistent with  $\rho_5 = 0.37$  and show that they imply very similar lifetime earnings persistence.

### 4.2.1 Efficiency Considerations

It is possible that  $\beta$  and  $\varphi$  could be estimated more efficiently by jointly estimating them with parameters of the earnings DGP.<sup>4</sup> Such an approach might also help with estimating the decomposition of  $Var(y(a_1))$  into  $Var(h)$  and  $Var(z(a_1)) + Var(\varepsilon)$ . Existing estimates of earnings DGPs often assume either that  $Var(h) = 0$  or that  $Var(z(a_1)) = 0$ . However, the joint estimation approach has the drawback of using small samples (in both time and cross-section dimension) of parent-child pairs for estimating the earnings DGP. Much better datasets are available, if the earnings DGP is estimated separately from the intergenerational persistence parameters.

A more promising approach is to estimate  $\beta$  and  $\varphi$  for a given DGP using a method of moments estimator based on the covariance of parent/child earnings at various ages. I do not pursue such estimators in this paper, mainly for reasons of comparability and transparency. Taking existing estimates of  $\rho_m$  as given ensures that my different conclusions about intergenerational persistence stem from the fact that lifetime earnings are not approximated well by the permanent component of earnings, and not from different ways of estimating  $\rho_m$ . Moreover, the results of the sensitivity analysis of section 5.1 suggest that different combinations of  $\beta$  and  $\varphi$  have very similar observable implications, so that identification with existing data would be weak. Nonetheless, future research should investigate more efficient estimation methods that fully exploit the restrictions implied by a given DGP.

## 4.3 Estimating Lifetime Earnings Persistence

Step 3 of the estimation approach computes the intergenerational persistence of lifetime earnings using a Monte Carlo simulation. For a given earnings DGP, 100 samples are drawn which contain earnings histories of 400 parents and their children. The size of each sample is typical for empirical studies and permits to calculate standard errors. For each individual,  $E_c$  is computed from its definition. Next,  $\rho_c$  is estimated from (1) using OLS. Finally, various estimates of  $\rho_m$  are computed according to methods proposed in the literature. Checking that the DGP replicates a variety of  $\rho_m$  estimates provides an indirect way of validating the DGP.

---

<sup>3</sup>It is interesting to note that the measure of lifetime earnings persistence commonly estimated in the literature ( $\beta$ ) is not even defined for the earnings DGPs most commonly estimated.

<sup>4</sup>Zimmerman (1992, section IV.D) jointly estimates an earnings DGP and  $\beta$ , conditional on  $\varphi = 0$ .

### 4.3.1 Lifetime Earnings Concepts

Estimating intergenerational persistence requires a precise definition of lifetime earnings. The literature generally agrees that the object of interest is some version of "average earnings received by an individual over his lifetime" (Mulligan 1997, p. 194), but does not provide an operational definition.<sup>5</sup> I therefore extract common definitions from studies that use estimates of intergenerational earnings persistence to parameterize general equilibrium models.

The most commonly used notion of lifetime earnings is the present discounted value of earnings over a person's lifetime defined as

$$E_P(i) = \sum_{a=a_1}^{a_R} Y(i, a) (1+r)^{a_1-a}$$

where  $r$  is a discount rate which I set to 0.04. Recall that  $Y(i, a)$  denotes realized earnings.  $E_P$  measures the contribution of labor earnings to a person's lifetime income. In its definition, zero earnings observations are retained. The age bounds should cover the person's entire lifetime, so that differences in labor force entry or exit ages contribute to earnings differences across individuals. This earnings concept is used, for example in studies of saving or intergenerational transfer behavior, such as Gokhale et al. (2001).

A second commonly used earnings concept is the present discounted value of full-time or potential earnings, which I denote  $E_F$ . It differs from  $E_P$  in that the present value is calculated for full-time earnings over a fixed age range that covers a typical work life. Dates with unusual work hours or unemployment spells are interpolated or workers with incomplete histories of full-time earnings are deleted.  $E_F$  is used to measure the distribution of lifetime labor endowments (e.g., Fullerton and Rogers 1993; Knowles 1999).

## 5 Findings

Table 4 shows the results. For each specification of the earnings DGP the table shows the persistence measures  $\rho_P$  and  $\rho_F$  for lifetime earnings and two estimates of  $\rho_1$  and  $\rho_5$ . Following most of the literature, the first estimate deletes observations with zero earnings. The second estimate follows Couch and Lillard (1998) in replacing zero earnings observations with one dollar.<sup>6</sup>

Consider first the findings for the baseline model based on STY's earnings process. The intergenerational persistence of  $h$  is chosen to match  $\rho_5 = 0.37$  for five-year average earnings with zero observations deleted. The model is also consistent with the findings of Solon 1992 and Zimmerman 1992 that single-year earnings persistence is roughly roughly 0.05 smaller than  $\rho_5$ . Three main insights emerge:

1. The permanent component of earnings ( $h$ ) is a poor proxy for lifetime earnings ( $E_P$  or  $E_F$ ). Lifetime earnings persistence, measured by  $\rho_P$  or  $\rho_F$ , is near 0.54, compared with  $\beta = 0.9$ .
2. Using a 5-year earnings average to proxy for lifetime earnings introduces a downward bias of one-third ( $\rho_5/\rho_P = 0.67$ ).
3. Retaining observations with zero earnings dramatically exacerbates this bias. Consistent with Couch and Lillard (1998),  $\rho_m$  drops to near 0.05 in that case.<sup>7</sup>

---

<sup>5</sup>The exception is Knowles (1999).

<sup>6</sup>In all cases, the standard errors across samples are close to 0.04.

<sup>7</sup>Retaining observations with at least one year of strictly positive earnings, Peters (1992) finds  $\rho_5 = 0.14$ . The model implies a very similar persistence of  $\rho_5 = 0.15$ .



[INSERT TABLE 4 HERE]

The intuition for these findings is straightforward. The difference between single-year and five-year average earnings is due to attenuation bias as is well understood from the discussion in Solon (1992) or Zimmerman (1992). Attenuation bias is also the reason for the difference between  $\rho_P$  and  $\rho_5$ . The reason why a five-year average is not sufficient to eliminate attenuation bias is the high persistence of the shocks to  $z$ . A parent with high earnings at age 45 could have experienced a positive shock at age 44 or at age 23. The data cannot identify the difference, except for the fact that age 23 shocks have "depreciated" somewhat, but not fully because of the high persistence.

Whether observations with zero earnings are excluded has little effect on the persistence of lifetime earnings. Zero earnings are essentially large transitory shocks, which do not contribute much to the present value of lifetime earnings (especially because individuals are observed relatively late in life). However, the effect on average earnings can be large. Since zero earnings observations are rare and not typical for a person's lifetime average earnings, it is appropriate to exclude them in the estimation of lifetime earnings persistence.

The model suggests that the persistence of average earnings depends *very strongly* on the age at which the parents are observed. At age  $a_1$  earnings are nearly as persistent as  $h$  because  $z(a_1, i) = 0$ . As parents get older, a larger fraction of the earnings variance is due to  $z$  which is not intergenerationally persistent. As a result, earnings persistence declines with parental age. This effect is quite strong as illustrated in Figure 1, which plots  $\rho_5$  against the average age at which parents are observed. For young parents,  $\rho_5$  is considerably larger than  $\rho_P$  and approaches the persistence of the permanent endowment ( $\beta = 0.9$  for the baseline DGP) as parental ages get closer to  $a_1$ . However, in the data, parents are typically observed past age 40. For example, the mean age of the fathers in the NLS sample of Zimmerman (1992) is 52 years; his estimates of  $\rho_4$  average 0.35. In Solon's (1992) PSID sample the average age is 44 and the resulting  $\rho_5$  is around 0.41. The simulation results are consistent with this age dependence of  $\rho_5$ . Observing the parents at ages 50 to 54 yields  $\rho_5 = 0.34$ , while observing parents at ages 42 to 46 yields  $\rho_5 = 0.38$ .<sup>8</sup> The age dependence of  $\rho_5$  is likely a robust feature of earnings processes where the transfer of earnings capacity takes place early in life.

## 5.1 Sensitivity Analysis

This section studies the robustness of the previous findings to alternative specification of the earnings process. A natural approach would be to study the implications of alternative DGPs estimated in the literature. However, for the DGPs summarized in table 2 it is not possible to generate enough intergenerational persistence. That is, these DGPs imply  $\rho_5 < 0.37$ , even if  $\varphi$  is set to the highest value consistent with  $Var(y(a_1)) = 0.3$ .<sup>9</sup> The reason is that the DGPs abstract from the permanent earnings component ( $h = 0$ ). Most of the variance of earnings at ages at which parents are observed in the data is therefore due to transitory shocks that are not intergenerationally persistent. Table 4 shows the results where  $\varphi$  is set to the largest feasible value. The qualitative conclusions of the baseline case are confirmed. In each case,  $\rho_5$  is below two-thirds of  $\rho_P$ . Not deleting observations with zero earnings widens the gap between  $\rho_5$  and lifetime earnings persistence.

It appears desirable to conduct a broader sensitivity analysis that varies parameters which are not estimated precisely over wide ranges. I take two features of the data to be uncontroversial. First, in cross-sectional data, the variance of log earnings rises from 0.3 at age 23 to 0.738 at age 50 (e.g., STY). Second, the persistence of the  $z$  process is larger than 0.9. I therefore consider STY's estimate  $\alpha = 0.977$  and  $\alpha = 0.9$ . Estimates of the variance of the iid shocks are more diverse. Fortunately, its exact value

<sup>8</sup>This experiment replicates the average age of parents in the data of Solon (1992) or Zimmerman (1992), but does not replicate the fact that their data contain parents within a wider age range.

<sup>9</sup>Given the value of  $Var(y(a_B))$  implied by the DGP, a higher value of  $\varphi$  would yield  $Var(y(a_1)) > 0.3$  even if  $\sigma_\omega = 0$ .

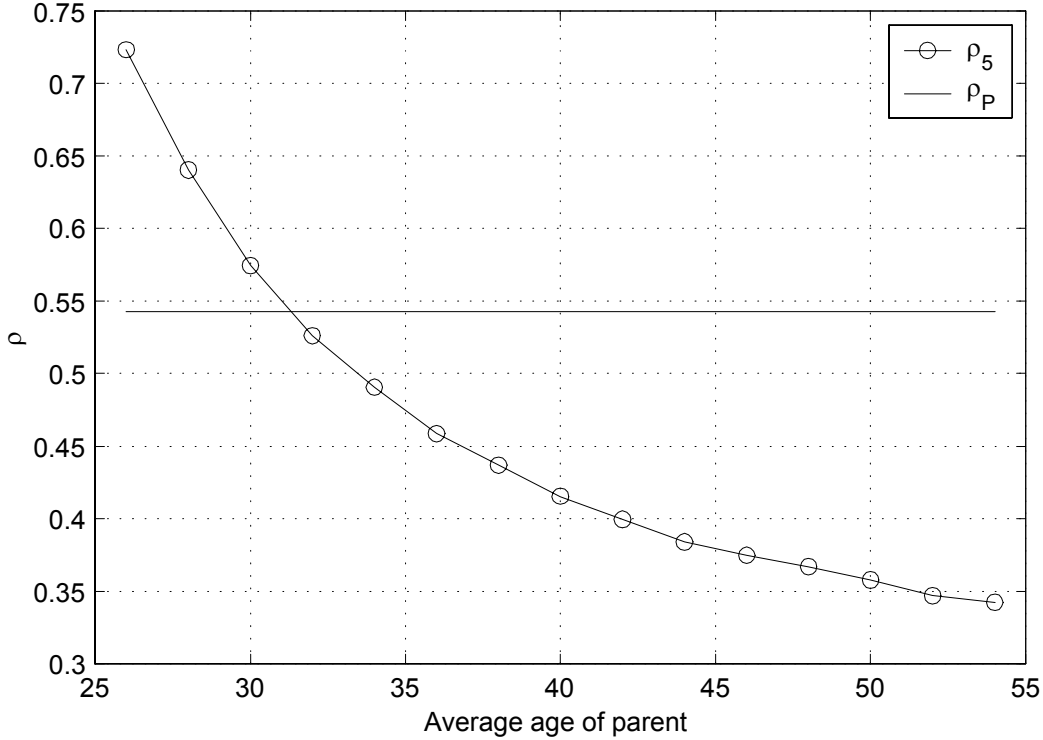


Figure 1: Parental age and intergenerational persistence

is not important for estimating intergenerational persistence. It mainly governs the difference between  $\rho_1$  and  $\rho_5$ . In what follows, I therefore set  $\sigma_\varepsilon^2 = 0.063$  based on STY.

Much less is known about the processes governing  $h$  and  $z(a_1)$ . In particular, it is not known which fraction of the variance of log earnings is due to  $h$  versus  $z(a_1)$ . Furthermore, it is not known which share of measured intergenerational earnings persistence is due to the persistence of  $h$  versus  $z(a_1)$ . The sensitivity analysis therefore covers the entire space of  $h$ -processes ( $\beta$  and  $\sigma_\eta^2$ ) that are consistent with measured intergenerational persistence and the variance of log earnings. For each  $h$ -process, the parameters governing  $z(a_1)$ , i.e.  $\varphi$  and  $\sigma_\omega^2$ , and the variance of the  $z$  shocks ( $\sigma_\zeta^2$ ) are chosen to match three observations:  $\rho_5 = 0.37$ , and the variance of log earnings at ages  $a_1$  and at age  $a^* = 50$ .

The algorithm involves the following steps.

1. Pick a  $\beta$  and the fraction of  $\text{Var}(y(a_1))$  attributed to  $h$ . Set  $\sigma_\eta^2 = \text{Var}(h)(1 - \beta^2)$  to match the stationary variance of  $h$ . Since the variance of  $z(a)$  must be positive,  $\text{Var}(h)$  is bounded above by the requirement  $\text{Var}(h) < \text{Var}(y(a)) - \text{Var}(\varepsilon)$  for all  $a$ .

2. Set the variance of the shock in the  $z$  process to match

$$\text{Var}(z(a^*)) = \text{Var}(y(a^*)) - \text{Var}(h) - \sigma_\varepsilon^2 = 0.738 \quad (4)$$

where  $\text{Var}(h)$  is fixed by step 1. Since  $\text{Var}(z(a))$  follows the difference equation

$$\text{Var}(z(a_1 + n)) = \text{Var}(\zeta) \left[ 1 + \alpha^2 + \dots + \alpha^{2(n-1)} \right] + \alpha^{2n} \text{Var}(z(a_1)) \quad (5)$$

the required shock variance is given by

$$\sigma_{\zeta}^2 = \frac{Var(z(a^*)) - \alpha^{2n} Var(z(a_1))}{1 + \alpha^2 + \dots + \alpha^{2(n-1)}}$$

3. Find the value of  $\varphi$  that matches  $\rho_5 = 0.37$ . This is not always possible, since  $\varphi$  is bounded above by the requirement that  $\sigma_{\omega}^2 > 0$ . Since,  $\sigma_{\omega}^2 = Var(z(a_1)) - \varphi^2 Var(z(a_B))$ , the upper bound for  $\varphi$  is given by

$$\varphi^2 \leq Var(z(a_1)) / Var(z(a_B)) \quad (6)$$

where  $Var(z(a_B))$  is calculated from (5).

### 5.1.1 Findings

The findings are shown in table 5. For the fraction of  $Var(y(a_1))$  that is due to  $h$  I consider the values 0.25, 0.5, and 0.75 (the values 0 and 1 are those of Huggett 1996 and Storesletten et al. 1998). For the intergenerational persistence of  $h$  is consider the values  $\beta = 0.9$  and 0.7 (lower values yield too small values for  $\rho_5$ ).

To illustrate the interpretation of table 5, consider the first case of the sensitivity analysis. The parameters of the  $h$ -process are fixed at  $\beta = 0.9$  and such that  $h$  accounts for 75% of  $Var(y(a_1))$ . Matching  $\rho_5 = 0.37$  then requires  $\varphi = 0.21$ . If zero earnings observations are not deleted,  $\rho_5$  falls to only 0.15. However, the intergenerational persistence of lifetime earnings is 0.54, regardless of whether zero earnings observations are interpolated ( $\rho_F$ ) or retained ( $\rho_P$ ). In the third case of table 5,  $Var(h)$  is reduced to 25% of  $Var(y(a_1))$ . It is then not possible to match  $\rho_5 = 0.37$ . Therefore,  $\varphi$  is set to its upper bound (6) which yields  $\rho_5 = 0.34$ .

The following key findings emerge from table 5. Consider first the cases with high  $z$  persistence ( $\alpha = 0.977$ ). In these cases it is possible to match  $\rho_5 = 0.37$ , but only if  $h$  accounts for a sufficiently large fraction of earnings variance and is sufficiently persistent. In all cases, using a 5-year earnings average to estimate the intergenerational persistence of lifetime earnings leads to a downward bias close to one-third ( $\rho_5$  approximately equals two-thirds of  $\rho_P$  or  $\rho_F$ ). Including observations with zero earnings implies persistence estimates close to those of Couch and Lillard (1998) which are much smaller than lifetime earnings persistence.

For the cases with lower persistence of  $z$  ( $\alpha = 0.9$ ) it is not possible to match the observed values of  $\rho_5$ . The downward bias of  $\rho_5$  is now even greater (40% or more). For the baseline DGP, the bias of  $\rho_P$  depended strongly on the age at which parents are observed. This remains true for processes where permanent shocks ( $h$ ) account for only a small part of earnings variance. For example, if  $h$  accounts for 25% of  $Var(y(a_1))$  and  $\beta = 0.9$ , then  $\rho_5 = 0.57$  if parents are observed around age 30, but falls to 0.3 for parents observed around age 55.

An important finding is that the bias due to proxying for lifetime earnings using a 5-year average is quite robust against variations in the DGP. This finding is illustrated in figure 2 which plots  $\rho_P$  against  $\rho_5$  for all cases of the sensitivity analysis, including the DGPs taken from the literature. Even if the deep parameters of the DGP cannot be precisely estimated, inference about lifetime earnings persistence remains possible.

A final concern is that  $z$  may be transmitted from parents to children at younger ages. The results presented so far assume that  $z$  is transmitted in the middle of childhood, which is at age 35, given that children typically co-reside with their parents between parental ages of 25 and 45. However, Heckman (1999) provides evidence that human capital endowments may be fixed early in a person's childhood. I therefore recalculate the sensitivity analysis for  $\alpha = 0.977$  under the assumption that  $z$  is transmitted at age  $a_B = 25$ . I find that it not possible to match  $\rho_5 = 0.37$  and that the gap between  $\rho_P$  and  $\rho_5$  is very similar to the baseline case of  $a_B = 35$  ( $\rho_P - \rho_5 \simeq 0.16$ ). I conclude that the findings of the

baseline case are robust against variations of the earnings process, at least within the class of processes that may be represented as a version of (3).

[INSERT TABLE 5 HERE]

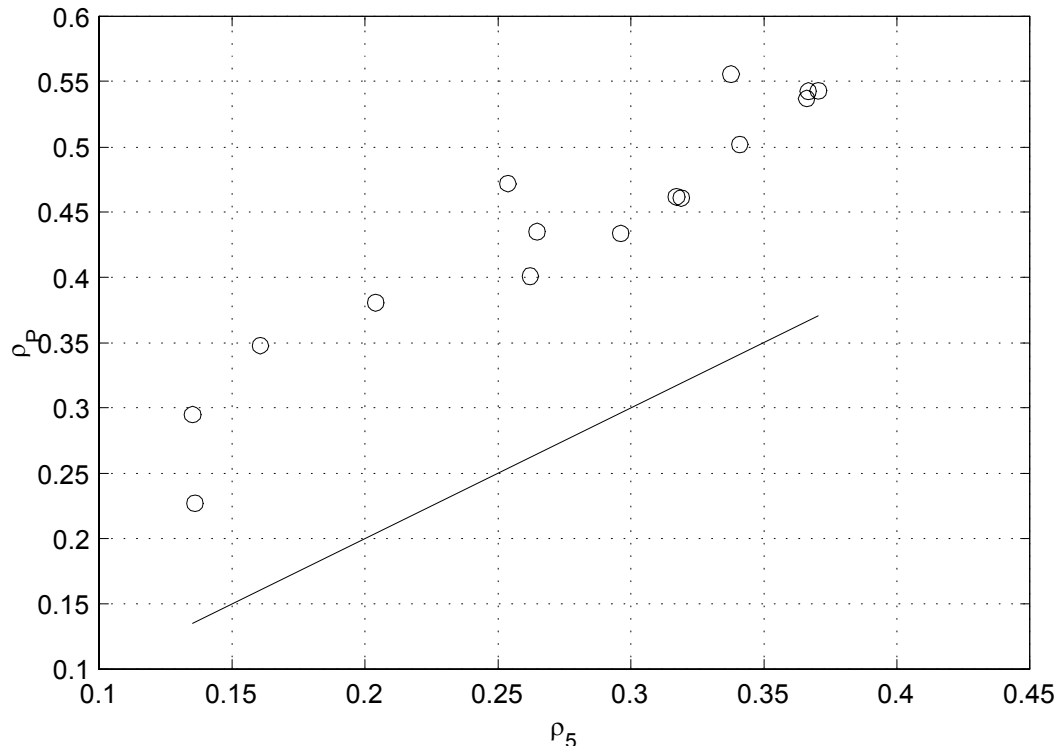


Figure 2: Sensitivity analysis

## 5.2 Alternative Proxies for Lifetime Earnings

A number of previous empirical studies have expressed concern with the use of average earnings as proxies for lifetime earnings. Below, I consider two alternative approaches based on fixed effect panel regressions and on instrumental variables.

### 5.2.1 Fixed Effect Estimators

One attempt at improving intergenerational persistence estimates based on average earnings is proposed by Knowles (1999) and Hendricks (2001). Both authors use direct estimates of the present value of lifetime earnings for  $E_c$  when estimating (1). However, while offering other benefits, their approach suffers from the same bias as the more common method which proxies for  $E_c$  with average earnings. In fact, if the DGP for earnings is described by (3), the two estimation methods are equivalent. Knowles and Hendricks estimate an individual fixed effect,  $y_0(i)$ , based on a regression of the form

$$y(i, a) - g(a) = y_0(i) + \tau(i)$$

over some fixed age range, where  $\tau(i)$  is a disturbance. Here I have assumed that the deterministic age profile,  $g(a)$ , is known. In practice, it is, of course, estimated jointly with the fixed effects. The OLS estimator of  $y_0(i)$  is the sample mean of  $y(i, a)$ . Intergenerational persistence is then estimated

according to (1), where the lifetime earnings concept is given by the present value

$$E_K(i) = \sum_{a=a_1}^{a_R} e^{y_K(i,a)} (1+r)^{a_1-a}$$

where  $y_K(i, a) = y(i, a)$  during ages where individual earnings are observed and  $y_K(i, a) = y_0(i) + g(a)$  otherwise. Since  $\ln E_K(i)$  is proportional to the fixed effect,  $y_0(i)$ , which in turn equals the sample mean of log earnings for person  $i$ , the resulting intergenerational persistence estimate equals  $\rho_5$ . This accounts for the fact that the persistence estimates of Knowles (1999) and Hendricks (2001) are similar to estimates of  $\rho_5$  published in the literature.<sup>10</sup>

### 5.2.2 Instrumental Variables

Instead of proxying for lifetime earnings using an earnings average, a number of studies rely on instrumental variable (IV) estimators. The instruments are chosen such that they are plausibly "correlated with father's permanent status, but uncorrelated with the transitory component of observed status" (Zimmerman 1992, p. 413). Examples include the father's education or occupation (Solon 1992; Mulligan 1997).

Perhaps the most fundamental problem with IV estimates of intergenerational persistence is that an IV estimator has a slightly different interpretation than the one applications typically have in mind. It may be paraphrased as: "If a parent earns  $x$  percent more than the mean *due to a particular attribute* (the instrumental variable), then the child earns, on average,  $\rho_{IV} x$  percent more than mean earnings." By contrast, applications typically consider the thought experiment: "Given a parent chosen *at random* from the population. If this parent earns  $x$  percent more than mean earnings, then his child earns, on average,  $\rho_c x$  percent more than mean earnings." This difference matters because the intergenerational persistence of earnings differences that are due to educational or occupational characteristics could be very different from that of earnings differences associated with other sources of earnings variation. The fact that estimates of  $\rho_{IV}$  are sensitive to the choice of instruments suggests that this may be a problem in practice.

A simple model of the instrumental variables approach is as follows. As an instrument for parental earnings a researcher uses a variable  $w(i)$  that is correlated with individual characteristics according to

$$\begin{aligned} w(i) &= \mu h(i) + (1 - \mu) z(a_{IV}, i) + \varepsilon_w \\ \varepsilon_w &\sim N(0, \sigma_w^2) \end{aligned} \tag{7}$$

Intergenerational persistence is estimated in two stages. The first stage regresses log earnings at age  $a_p$  on the instrument

$$y(a_p, i) = \xi_0 + \xi_1 w(i) + \xi_i^*$$

The second stage regresses child earnings on predicted parental earnings:

$$y(a_c, i) = \theta + \rho_{IV} \widehat{y}(a_p, p(i)) + v_i$$

where predicted earnings are given by  $\widehat{y}(a_p, i) = \widehat{\xi}_0 + \widehat{\xi}_1 w(i)$ . This is a fairly general representation of IV approaches in the literature. Essentially, different instruments differ in the parameters of equation (7),  $\mu$  and  $a_{IV}$ .

If the objective is to estimate the intergenerational persistence of permanent earnings, IV is a viable approach. If an instrument can be found which is correlated with  $h$  but not with  $z$  ( $\mu = 1$ ), then  $\rho_{IV}$

---

<sup>10</sup>However, their approach has other benefits compared with the more common approach using average earnings. It permits to estimate separate age profiles,  $g(a)$ , for different demographic groups (Knowles 1999). It also permits to use all available earnings observations for an individual, instead of a fixed number  $m$  (Hendricks 2001).

consistently estimates  $\beta$ . However, if the objective is to estimate the intergenerational persistence of lifetime earnings ( $\rho_P$  or  $\rho_F$ ), instrumental variable methods encounter additional problems. Lifetime earnings persistence is a complicated function of all the parameters characterizing the DGP. Consistently estimating  $\rho_c$  then requires an instrument that attaches the correct weights to  $h$  and  $z$ . But these weights are, of course, unknown.

Without information about the weight ( $\mu$ ) associated with a particular instrument, little can be said about the relationship between  $\rho_{IV}$  and parameters of interest. It is not even ensured that  $\rho_{IV}$  will lie between the persistence coefficients of  $h$  and  $z$ . To illustrate this point, consider the following example. Assume that the earnings DGP is given by the first row of the sensitivity analysis in table 4. The parameters of equation (7) are given by  $\sigma_w^2 = 0.2$  and  $a_{IV} = 45$ . Parents are observed at age  $a_p = 47$  which is the mid point of the age interval used for averaging. In this example, an instrument that is only correlated with  $z$  ( $\mu = 0$ ) yields  $\rho_{IV} = 0.29$ , even though both  $z$  and  $h$  are highly persistent ( $\varphi = 0.62$  and  $\beta = 0.9$ ). The intuition is that the correlation between  $z(a_1, i)$  and  $z(a_{IV}, i)$  is not that large. Hence,  $z(a_{IV}, i)$  is far less persistent across generations than is  $z(a_1, i)$ .

Of course, these numbers are mere examples. Without information on how the instruments relate to the individual characteristics that determine earnings, the relationship between  $\rho_{IV}$  and  $\rho_c$  cannot be determined. As a result, it is difficult to infer lifetime earnings persistence from instrumental variable methods. They may, however, be useful for gaining insight into the DGP for earnings. To the extent that instruments can be found that are correlated with  $h$ , but not with  $z$  (or vice versa), these can be used to estimate  $\beta$  (or  $\varphi$ ). This information can then be used in Monte Carlo simulations to help infer  $\rho_c$ .

### 5.2.3 Group Estimators

A particular type of instrumental variables estimator could, in principle, overcome some of the estimation problems described above. The idea is to sort observations into groups that differ in their permanent earnings ( $h$ ), but not in their transitory earnings ( $z$  and  $\varepsilon$ ).<sup>11</sup>

The details are as follows. Decompose log lifetime earnings into a permanent and a transitory component:

$$\ln E_P(i) = h(i) + \mu(i)$$

where, from the definition of  $E_P$ ,

$$\mu(i) = \ln \left( \sum_{a=a_1}^{a_R} \exp [z(i, a) + \varepsilon(i, a)] R_a \right)$$

and  $R_a$  is a discount factor for age  $a$ . Actually observed is an  $m$  year earnings average starting at age  $a_A$ :

$$\begin{aligned} \ln E_m(i) &= h(i) + \mu_m(i) \\ \mu_m(i) &= \sum_{a=a_A}^{a_A+m-1} [z(i, a) + \varepsilon(i, a)] / m \end{aligned}$$

Note that

$$\ln E_m(i) = \ln E_P(i) + \Delta\mu(i)$$

where  $\Delta\mu(i) \equiv \mu_m(i) - \mu(i)$ . If  $\Delta\mu(i)$  has mean zero, then average earnings are a valid proxy for lifetime earnings. Of course, this does not mean that using  $E_m(i)$  in (1) yields a consistent estimate of

<sup>11</sup>The group estimation approach has, to my knowledge, not been implemented in the literature. I am grateful to Seung Ahn for suggesting this approach.

$\rho_P$ . This would require in addition that for each individual  $\Delta\mu(i) = 0$ . However, it provides the basis for constructing a group estimator.

If it possible to group observations such that within each group the  $\Delta\mu(i)$  average to the same value, but the  $h(i)$  averages differ, then the group mean of  $\ln E_m$  equals that of  $\ln E_P$ . As a result, OLS consistently estimates  $\rho_P$  once individual values of  $E_m$  are replaced by their group averages. An intercept is needed in (1), if the mean of  $\Delta\mu(i)$  is not zero. This approach has two benefits over the one proposed in this paper: it requires weaker assumptions on the earnings DGP and it has known econometric properties (e.g., standard errors).

The key difficulty is how to define the groups. If (1) were a structural equation, then any grouping scheme that generates differences in average  $h$  across groups, but maintains the same averages for  $\Delta\mu(i)$  would be acceptable. Examples include some of the instruments used for parental earnings in the literature. Mulligan (1997) uses average earnings of parents in the same education, occupation, or industry class as instruments. Solon (1992) uses the father's education.<sup>12</sup>

One potential problem is that large samples may be required to ensure that the group averages of  $\Delta\mu(i)$  are close to each other. A second problem is that  $z$  may be intergenerationally persistent. Perhaps a more fundamental problem, which the group estimator shares with IV approaches, is that a group estimator reveals the intergenerational persistence of earnings variations that are due to the grouping characteristic (e.g., education). By contrast, the object of interest in applications is the intergenerational persistence of earnings for randomly chosen parents. I conclude that group estimators are feasible in principle, but will likely encounter practical problems.

## 6 Conclusion

This paper proposes a new method for estimating the intergenerational persistence of lifetime earnings from data that contain only short sections of individual earnings histories. I find that lifetime earnings are substantially more persistent than estimates of average earnings persistence suggest. The coefficient in a regression of children's lifetime earnings on fathers' lifetime earnings is approximately 0.54. Proxying for lifetime earnings using five year averages leads to a downward bias in estimated intergenerational persistence of one-third. The bias is much stronger, if observations with zero earnings are not excluded from the sample. These findings are robust against alternative assumptions about the data generating process for earnings.

Future research should apply the methods developed in this paper to study the intergenerational persistence of other indicators of economic status, such as income, wage rates, and consumption.

---

<sup>12</sup>Solon (1992) also discusses an additional source of bias that might apply to other instruments and to the group estimator: parental education could have a direct effect on child earnings.

## References

- [1] Behrman, Jere R.; Paul Taubman (1985). "Intergenerational Earnings Mobility in the United States: Some Estimates and a Test of Becker's Intergenerational Endowments Model." *Review of Economics and Statistics* 67(1): 144-51.
- [2] Castaneda, Ana; Javier Diaz-Gimenez; Jose-Victor Rios-Rull (2000). "Accounting for earnings and wealth inequality." Mimeo. University of Pennsylvania.
- [3] Couch, Kenneth A.; Dean R. Lillard (1998). "Sample selection rules and the intergenerational correlation of earnings" *Labour Economics* 5(3): 313-29.
- [4] Fullerton, Don; Diane L. Rogers (1993). *Who Bears the Lifetime Tax Burden?* Brookings: Washington, DC.
- [5] Gokhale, Jagadeesh; Laurence J. Kotlikoff; James Sefton; Martin Weale (2001). "Simulating the transmission of wealth inequality via bequests." *Journal of Public Economics* 79: 93-128.
- [6] Gourinchas, Pierre-Olivier (2000). "Precautionary savings, lifecycle and macroeconomics." Mimeo. Princeton University.
- [7] Heckman, James J. (1999). "Policies to foster human capital." NBER working paper #7288.
- [8] Hendricks, Lutz (2001). "How Do Taxes Affect Human Capital? The Role of Intergenerational Mobility." *Review of Economic Dynamics* 4(3): 695-735.
- [9] Hubbard, Glenn; Jonathan Skinner; Steven Zeldes (1995). "Precautionary saving and social insurance." *Journal of Political Economy* 103(2): 360-99.
- [10] Huggett, Mark (1996). "Wealth distribution in life-cycle economies." *Journal of Monetary Economics* 38: 469-94.
- [11] Knowles, John (1999). "Measuring Lifetime Inequality." Mimeo. University of Pennsylvania.
- [12] Mulligan, Casey B. (1997). *Parental priorities*. Chicago: University of Chicago Press.
- [13] Peters, H. Elizabeth (1992). "Patterns of intergenerational mobility of income and earnings." *Review of Economics and Statistics* 74(3): 456-66.
- [14] Solon, Gary (1992). "Intergenerational Mobility in the United States." *American Economic Review*, 82(3), 393-408.
- [15] Storesletten, Kjetil; Chris Telmer; Amir Yaron (1998). "Consumption and risk sharing over the life cycle." Mimeo. Stockholm University.
- [16] Zimmerman, David J. (1992). "Regression toward mediocrity in economic stature." *American Economic Review*, 82(3): 409-29.



Table 1: Empirical estimates of intergenerational earnings persistence

Source	Data source	$\rho_1$	$\rho_2$	$\rho_3$	$\rho_4$	$\rho_5$	Avg age Fathers	Avg age Sons	Sample selection
Mulligan (1997)	PSID				0.320		40	>30	Earnings > 0
Solon (1992)	PSID	0.353	0.384	0.403	0.409	0.413	44	29	Earnings > 0
Couch and Lillard (1998)	PSID	0.432	0.502	0.534	0.524	0.531	42	29	Earnings > 0
	PSID	-0.019	-0.018	-0.017	-0.018	-0.019	42	29	All individuals included
	NLS	0.295	0.325	0.348	0.372		49	32	As Zimmerman (1992)
	NLS	0.096	0.102	0.112	0.117		49	32	All individuals included
Peters (1992)	NLS					0.140	47	30	Earnings > 0 in at least 1 year
Zimmerman (1992)	NLS	0.294			0.346		52	36	Earnings > \$3000 in 1984 dollars. At least 30 hours / week. At least 30 weeks / year

Table 2: Empirical estimates of earnings processes

Source	Var( $h$ )	$\alpha$	Var( $\zeta$ )	Var( $z(a_1)$ )	Var( $\varepsilon$ )
Storesletten et al. (1998)					
Exactly identified	0.242	0.984	0.022	0.000	0.057
Overidentified	0.244	0.977	0.024	0.000	0.063
Huggett (1996)		0.960	0.071	0.380	
Hubbard, Skinner, Zeldes (1995)					
Less than high school		0.955	0.033		0.040
At least high school		0.946	0.025		0.021
At least college degree		0.955	0.016		0.014
Gourinchas (2000)		1.000	0.021	0.021	

Notes: Var( $z(a_1)$ ) is not reported by Hubbard, Skinner, and Zeldes (1995).

Table 3: Baseline parameters

Demographics:	$a_1 = 23; a_R = 65; a_B = 35$
$h$ process:	$\beta = 0.9; \sigma_\eta = 0.214$
$z$ process:	$\alpha = 0.977; \sigma_\zeta = 0.155$
$z_1$ process:	$\varphi = 0; \sigma_\omega = 0$
iid shocks:	$\sigma_\varepsilon = 0.251$
Other:	$r = 0.04$

Table 4: Simulation results

	$\beta$	Var( $h$ )	$\varphi$	$\rho_A$	$\rho_B$	Zeros deleted		Zeros retained	
						$\rho_1$	$\rho_5$	$\rho_1$	$\rho_5$
Data						0.32	0.37	0.04	0.05
Baseline	0.90	80.7	0.00	0.54	0.54	0.33	0.37	0.05	0.15
Huggett (1996)	n/a	n/a	0.73	0.40	0.39	0.29	0.26	0.06	0.12
Gourinchas (2000)	n/a	n/a	0.28	0.23	0.22	0.14	0.14	0.02	0.05

Notes: The table shows the means over 100 samples, each containing 400 parent/child pairs.

Var( $h$ ) denotes the fraction of Var( $y(a_1)$ ) that is due to  $h$ .

Table 5: Sensitivity analysis

	$\beta$	Var( $h$ )	$\varphi$	$\rho_A$	$\rho_B$	Zeros deleted		Zeros retained	
						$\rho_1$	$\rho_5$	$\rho_1$	$\rho_5$
$\alpha = 0.977$	0.9	75	0.21	0.54	0.54	0.34	0.37	0.06	0.15
	0.9	50	0.50	0.54	0.53	0.35	0.37	0.06	0.15
	0.9	25	0.62	0.50	0.50	0.33	0.34	0.06	0.14
	0.7	75	0.21	0.43	0.43	0.27	0.30	0.05	0.12
	0.7	50	0.50	0.46	0.46	0.30	0.32	0.05	0.13
	0.7	25	0.62	0.46	0.46	0.31	0.32	0.05	0.13
$\alpha = 0.900$	0.9	75	0.16	0.56	0.56	0.28	0.34	0.05	0.14
	0.9	50	0.41	0.47	0.47	0.21	0.25	0.04	0.10
	0.9	25	0.53	0.35	0.35	0.14	0.16	0.03	0.06
	0.7	75	0.16	0.43	0.44	0.22	0.26	0.04	0.11
	0.7	50	0.41	0.38	0.38	0.17	0.20	0.03	0.08
	0.7	25	0.53	0.30	0.29	0.12	0.14	0.02	0.05

Notes: Var( $h$ ) denotes the fraction of Var( $y(a_1)$ ) that is due to  $h$ .