The Intergenerational Persistence of Lifetime Earnings*

Lutz Hendricks
University of Sydney, Iowa State University, CESifo, and Center for Financial Studies, Frankfurt
October 17, 2005

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 previous estimates based on short panel data suggest. About 54% of lifetime earnings differences between fathers persist into their sons' generation. This persistence estimate exceeds previous estimates based on five year earnings averages by one third. These findings are robust against alternative assumptions about the data generating process for earnings.

Keywords: Intergenerational persistence. Intergenerational mobility. JEL: J62

1 Introduction

This paper studies the measurement of intergenerational earnings mobility. The specific question addressed is: To what extent do differences in lifetime earnings persist from parents to children? A sizeable literature has addressed this question by estimating reduced form earnings equations of the form

$$\ln E_c(i) = \theta + \rho_c \ln E_c(p(i)) + \upsilon(i) \tag{1}$$

where $E_c(i)$ denotes child i's lifetime earnings, $p(i)$ is child i's parent, and $\upsilon(i)$ is a random shock of mean zero.[1] The coefficient ρ_c measures the fraction of parental lifetime earnings gaps that is, on average, preserved among the children. The value of ρ_c is important as a

---

*I thank the editor, two anonymous referees, and seminar participants at IZA (Bonn) for helpful comments. Seung Ahn provided invaluable advice on econometric issues.

[1] The survey by Solon (1999) summarizes this literature.

descriptive statistic of intergenerational mobility. It is also an important component of the earnings processes of computable general equilibrium models (cf. Castaneda et al., 2003).[2]

In spite of numerous efforts to estimate ρ_c, no consensus has emerged as to whether intergenerational persistence is weak or strong. Recent estimates of ρ_c for fathers and sons range from almost complete mobility ($\rho_c = 0.05$ in Couch and Lillard, 1998) to strong persistence ($\rho_c = 0.6$ in Mazumder, 2003) with the majority of the estimates clustering around 0.4 (Solon, 1999). 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 ρ_c is that (1) should be estimated for *lifetime* earnings, which are defined as the discounted present value of earnings over 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 directly observable. The literature has addressed this problem by using various proxies for lifetime earnings. The most common proxies are earnings averages over up to five years or instrumental variables (Zimmerman, 1992; Solon, 1992). Unfortunately, different proxies yield different estimates of intergenerational persistence (see the discussion of section 2). This raises the question how lifetime earnings persistence can be inferred from available data.

To address this problem, this paper proposes a new estimation approach which recognizes that lifetime earnings persistence is a function of the stochastic process governing earnings over the life-cycle. The first step of the proposed method is therefore to model this process and to estimate its parameters. Lifetime earnings persistence is then computed from simulated earnings histories for a large number of persons. One complication is that no single dataset contains both long earnings histories suitable for estimating the properties of earnings shocks and observations for parents and their children. It is therefore advantageous to break the estimation of the earnings process into two steps. Step 1 estimates the parameters not related to intergenerational persistence, such as the variance and persistence of earnings shocks. Many such estimates have been reported in the literature based on panel data with a long time dimension. Step 2 estimates the intergenerational persistence parameters using a method of moments approach. Essentially, these parameters are chosen to match the intergenerational persistence of short earnings averages that can be observed datasets of fathers and sons.

This approach has several benefits. (*i*) The information from various estimation methods and samples can be combined to arrive at the final estimate of ρ_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 instruments provide good proxies for lifetime earnings. No such proxies are needed here because the relationship between observed objects (e.g., earnings aver-

---

[2] Another notion of intergenerational persistence measures how an exogenous change in parental earnings affects child earnings. Since equation (1) is not a structural relationship, the value of ρ_c sheds little light on this.

ages) and lifetime earnings is modeled explicitly. (*iii*) The estimate of ρ_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 data generating process (DGP) for earnings. However, other methods of estimating intergenerational persistence share this drawback; it is merely not as apparent if the DGP is not explicitly stated.[3] Fortunately, the findings turn out to be 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.

Perhaps the main contribution of the proposed method is *clarity.* The objects to be estimated (lifetime earnings and its persistence) are clearly defined. How lifetime earnings persistence relates to the deep parameters of the DGP and how the deep parameters relate to observable data is explicitly stated. A large part of the literature consists of suggestions on how to estimate lifetime earnings persistence from imperfect data with minimal bias. The point of this paper is to argue that the bias cannot be understood without explicitly defining the DGP and the concept of lifetime earnings to be measured.[4]

I consider two definitions of lifetime earnings that are commonly used in computable general equilibrium models: the discounted present value of realized lifetime earnings (E_P) and the discounted present value of full-time or potential earnings (E_F).[5] 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

---

[3] The exception is Haider and Solon (2003) where the sample is large enough and long enough to estimate the relationship between average earnings and lifetime earnings nonparametrically. Mazumder (2003) also recognizes the importance of specifying an earnings DGP for inferring the deep intergenerational persistence parameters from estimates of ρ_m. He uses a large dataset of Social Security earnings histories to estimate an earnings model that is far more general than the class of models used in this paper. However, instead of estimating lifetime earnings persistence, Mazumder estimates the persistence of the permanent component of the earnings DGP (the parameter β in equation 4).

[4] Readers may disagree with my definition of lifetime earnings or with my assumptions about the DGP. This is not a fundamental problem. The method proposed here permits to calculate persistence for any definition of lifetime earnings and for a rich class of DGPs from data containing only short sections of individual earnings histories.

[5] See section 3 for references to studies which use these definitions.

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. My preferred estimate of intergenerational persistence of lifetime earnings for fathers and sons is 0.54 for both E_P and E_F. This compares with previous estimates around 0.37 based on studies that use five year earnings averages as proxies for lifetime earnings (section 2 summarizes these studies). One reason why average earnings are less persistent than lifetime earnings is attenuation bias. Empirical estimates of the stochastic process governing earnings indicate that transitory shocks account for a large part of the dispersion of earnings around middle age, which is when fathers are typically observed. Average earnings are therefore a noisy estimate of lifetime earnings. Measurement error in the regressor $E_c(p(i))$ in (1) then leads to attenuation bias (Solon, 1992; Zimmerman, 1992).

Couch and Lillard (1998) and Peters (1992) suggest that years with zero earnings should be retained when computing average earnings. This greatly exacerbates attenuation bias. Empirically plausible earnings processes imply intergenerational persistence of five year earnings averages around 0.1 when zero earnings are retained. Yet lifetime earnings persistence is much higher (0.54). The intuition is that very low or zero earnings observations are rare. They make a small contribution towards lifetime earnings, but have a strong effect on short earnings averages.

Some studies have proposed instrumental variable estimators as a way of separating permanent from transitory earnings components. I argue in section 5.2.2 that such estimators are useful for estimating the persistence of the permanent component of earnings, but less useful for estimating the persistence of lifetime earnings as measured by E_P or E_F.

The findings reported in this paper are obtained from a class of earnings processes that are commonly used in the macroeconomics literature. The findings are robust within this class. However, more complicated earnings processes have been proposed in the labor literature. These encompass age-dependent shock variances and heterogeneity in mean age-earnings profiles (e.g., Baker and Solon, 2003). It would be straightforward to extend the method proposed here to these processes. This is, in fact, an important benefit of the approach: it is possible to accommodate a wide range of earnings DGPs and to incorporate information obtained from multiple studies or datasets.

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 estimates of intergenerational earnings persistence which are summarized in Table 1.[6] It is helpful to establish notation first. I denote

---

[6] For comprehensive surveys see Mulligan 1997 and Solon 1999.

an m year average of log earnings by E_m and its intergenerational persistence by ρ_m. Instrumental variable estimates of lifetime earnings are denoted E_{IV} with a persistence coefficient of ρ_{IV}.

Early studies, such as Behrman and Taubman (1985), estimated (1) based on earnings in a single year, resulting in estimates of ρ_1 near 0.25. The innovation of the more recent literature was to recognize that OLS estimates suffer from attenuation bias (Behrman and Taubman, 1990, Solon, 1992, Zimmerman, 1992). Using more representative samples and proxying for lifetime earnings with average earnings over up to five years, these studies arrive at persistence estimates on the order of 0.35 to 0.4. Using average earnings over up to 16 years, Mazumder (2001) estimates ρ_{16} near 0.6. Couch and Lillard (1998) and Peters (1992), on the other hand, argue that these higher estimates are biased *upward* because of restrictive sample selection criteria, especially the deletion of observations with zero or low earnings. If such observations are included, much lower estimates of ρ_m, typically below 0.1, are obtained.

Instrumental variable methods avoid the use of average earnings as a proxy for lifetime earnings. Using variables such as education or occupation as instruments for permanent earnings yields persistence estimates that are somewhat larger than those obtained from earnings averages. However, the findings are sensitive to the choice of instruments, ranging from 0.36 (Zimmerman, 1992) to 0.53 (Solon, 1992).

This selective review shows that previous studies differ in the 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 ρ_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. Finally, not deleting observations with zero earnings reduces ρ_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.

In what follows, these estimates of ρ_m will form the basis for estimating lifetime earnings persistence. It is important to note, however, that the estimation method proposed in this paper is more general. One of its important benefits is the ability to incorporate information from a variety of studies and data sources about the intergenerational persistence of observable lifetime earnings proxies. For example, a more complete analysis should take into account how the persistence of average earnings varies with the ages at which parents and children are observed.[7]

[INSERT TABLE 1 HERE]

---

[7] The literature calls this "life-cycle bias." See Solon (1999); Mazumder (2001); Grawe (2002).

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

$$y(i,a) = g(a) + h(i) + z(i,a) + \varepsilon(i,a) \tag{2}$$

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: (i) A deterministic age profile, $g(a)$. (ii) A permanent component $h(i)$ with intergenerational persistence β. (iii) A persistent component, $z(i,a)$, which evolves as an AR(1) process with persistence α. (iv) 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. Since persons enter the model as young adults, it appears reasonable to assume most parental actions that could influence child earnings take place before age a_1.[8]

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 $y(i,a)$.

The motivation for restricting the stochastic process for earnings to the functional form (2) is that such processes are frequently used in applied work. In particular, a large number of computable general equilibrium models employ versions of (2). Table 2 displays some examples, although many more could be listed.

Since (2) is a reduced form model, the value of ρ_c is not useful for identifying the causal effect of parental earnings on child earnings or for drawing policy conclusions; a structural model of parental investments in children is needed for that.[9] The value of ρ_c is, however,

---

[8] Heckman (1999) provides evidence of the importance of early human capital formation.

[9] Rubinstein and Tsiddon (2004) develop such a model and offer some evidence that the intergenerational persistence of earnings and education has changed over time.

an important descriptive measure of intergenerational mobility (see Solon, 1999 and the references therein). It is also a necessary input for the specification of earnings processes in computable general equilibrium models (e.g., Fullerton and Rogers, 1993; Castaneda et al., 2003; Hendricks, 2001).

[INSERT TABLE 2 HERE]

4 Estimation

The method for estimating the earnings model (2) is motivated by the fact that no single dataset contains both long earnings histories and observations for fathers and sons. It is therefore advantageous to estimate the parameters that govern earnings over the life-cycle separately from those governing intergenerational persistence. The estimation thus proceeds in three steps, which are described in detail in the following three subsections.

1. Estimate the parameters of the earnings DGP (2) not related to intergenerational persistence from long panel data.
2. Estimate the parameters governing intergenerational persistence (β and φ) from observations for fathers and sons.
3. Estimate lifetime earnings persistence by simulating earnings histories for fathers and sons.

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 (β and φ) are estimated using a 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 φ becomes irrelevant. The other DGPs set $h = 0$, so that β becomes irrelevant.[10] Then a single parameter, either β or φ, 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 $a_B = 35$, 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 β and φ by matching two or more intergenerational persistence observations. However, the sensitivity analysis shows that alternative combinations of β and φ that yield the same value of ρ_5 are nearly observationally equivalent. As a result, identification would be weak. I therefore explore all possible combinations of β and φ that are consistent with $\rho_5 = 0.37$ and show that they imply very similar lifetime earnings persistence.

A promising extension is to estimate β and φ 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 ρ_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 ρ_m. Moreover, the results of the sensitivity analysis of section 5.1 suggest that different combinations of β and φ have very similar observable implications, so that identification with existing data would be weak.

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. At the

---

[10] This leads to an interesting contradiction. Many studies use β as their measure of intergenerational persistence. Yet β is meaningless for the DGPs specified in many applied papers, which set $h = 0$.

same time, averaging over all samples yields precise point estimates. For each individual, E_c is computed from its definition. Next, ρ_c is estimated from (1) using OLS. Finally, various estimates of ρ_m are computed according to methods proposed in the literature. Checking that the DGP replicates a variety of ρ_m estimates provides an indirect way of validating the DGP.

In computable general equilibrium models, two concepts of *lifetime earnings* are used repeatedly. The first concept, E_P, 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, 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).

Note that the estimation method proposed here is insensitive to serially uncorrelated measurement error in the earnings data. Measurement error implies that the variance of the disturbance in the earnings DGP (2) is overstated. The estimation algorithm then simulates individual earnings including measurement error. It matches the observed persistence of average earnings (ρ_m), again including measurement error. As a result, the estimate of ρ_m will be biased. However, to the extent that serially uncorrelated measurement error averages out in the simulation of E_P or E_F, its effect on the estimated persistence of lifetime earnings is minimal.

4.4 Discussion

The innovation of the proposed method is to exploit information about the stochastic process governing earnings in order to estimate lifetime earnings persistence. Due to data limitations, previous studies used short earnings averages as proxies for permanent earnings. A common approach decomposes log earnings at age a into a permanent component (h) and a transitory component (z):[11]

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

---

[11] 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.

The permanent component is transmitted from parents to children according to

$$h(i) = \beta\, h(p(i)) + \eta(i) \tag{4}$$

where $\eta(i)$ is a random shock. Equation (4) is then estimated using a proxy for h, such as average earnings over five years or instrumental variables. The resulting estimate of β is interpreted as lifetime earnings persistence.[12] Two important assumptions underlie this approach. (i) A good proxy for h can be found. (ii) The intergenerational persistence of h is close to that of lifetime earnings.

What motivates the use of h as a proxy for permanent earnings is the intuition that 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.[13]

The findings reported below show that, for plausible earnings processes, the resulting estimates of β can be very different from lifetime earnings persistence. Therefore, the estimation method proposed here avoids equating permanent earnings (h) with lifetime earnings (E_c). Instead, it explicitly models lifetime earnings persistence as a function of the earnings DGP.

5 Findings

Table 4 shows the results. For each specification of the earnings DGP the table shows the persistence measures ρ_P and ρ_F for lifetime earnings and two estimates of ρ_1 and ρ_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.[14]

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 ρ_5. Three main insights emerge:

---

[12] 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.

[13] Since this paper was written, Haider and Solon (2003) have made the related point that the error correction model commonly used in the literature (cf. equation 3) misspecifies the relationship between current and lifetime earnings.

[14] In all cases, the standard errors across samples are close to 0.04.

1. The permanent component of earnings (h) is a poor proxy for lifetime earnings (E_P or E_F). Lifetime earnings persistence, measured by ρ_P or ρ_F, is near 0.54, compared with $\beta = 0.9$.
2. Using a five-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), ρ_5 drops to 0.15 in that case.[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 ρ_P and ρ_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 effect on lifetime earnings would differ substantially. However, 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.[16]

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.[17]

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 ρ_5 against the average age at which parents are observed. For young parents, ρ_5 is considerably larger than ρ_P and approaches the persistence of the permanent endowment ($\beta = 0.9$ for the baseline DGP) as parental ages get closer to a_1. The age dependence of

---

[15] 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 $\rho_5 = 0.15$.

[16] One reader expressed concerns that such highly persistent shocks to z would be indistinguishable from permanent shocks to h in the data. The sensitivity analysis shows that there is no need to empirically distinguish such shocks. Stochastic processes that are consistent with the key features of the data described in section 5.1 yield similar lifetime earnings persistence.

[17] This conclusion could change, if the frequency of zero earnings is higher for less able workers. A structural model of hours choice is needed to assess this possibility.

ρ_5 is likely a robust feature of earnings processes where the transfer of earnings capacity takes place early in life.

[Insert figure 1 here]

A similar age dependence is found in empirical studies. Grawe (2002) documents that studies which observe parents at later ages tend to estimate lower values of ρ_5. Given that these studies differ in many other respects, it is difficult to draw quantitative conclusions from them. However, the predictions of my model are roughly consistent with the findings of two prominent studies. Zimmerman (1992) obtains estimates of ρ_4 around 0.35 from an NLS sample where fathers are on average 52 years old. In Solon's (1992) PSID sample, the average age of the fathers is 44 years and the resulting ρ_5 is around 0.41. In the model, observing the parents at ages 42 to 46 yields $\rho_5 = 0.38$, while observing parents at ages 50 to 54 yields $\rho_5 = 0.34$.[18]

The strong age dependence of ρ also helps reconcile the diverse, though limited, evidence based on longer earnings averages. Based on PSID parents observed for at least 10 years at ages up to 65, Hendricks (2001) estimates intergenerational persistence coefficients around 0.3. For parents observed between the ages of 45 and 54, the baseline model implies $\rho_{10} = 0.23$. This is somewhat lower than the finding of Hendricks (2001). In part, the difference may be due to the fact that the PSID sample contains parents of diverse ages. Mazumder (2001) estimates ρ_{16} around 0.6 based on Social Security earnings of parents with average ages between 33 and 48. The baseline model implies $\rho_{16} = 0.44$ when parents are observed between these ages. The model suggests that a large part of the large gap between the findings of Hendricks and Mazumder may be due to the differences in parental ages.

Comparing the empirical findings with the model predictions, it appears that the baseline model underpredicts the persistence of long earnings averages. However, given the limited evidence and the selection problems of Social Security data, more work is needed before strong conclusions can be drawn. This suggests a promising avenue for future research. The strong dependency of ρ_m on m and on parental age could be exploited to obtain further insights into which components of the earnings process are transmitted from parents to children.

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 φ is set to the highest value consistent with $Var(y(a_1)) = 0.3$.[19] The

---

[18] 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.

[19] Given the value of $Var(y(a_B))$ implied by the DGP, a higher value of φ would yield $Var(y(a_1)) > 0.3$ even if $\sigma_\omega = 0$.

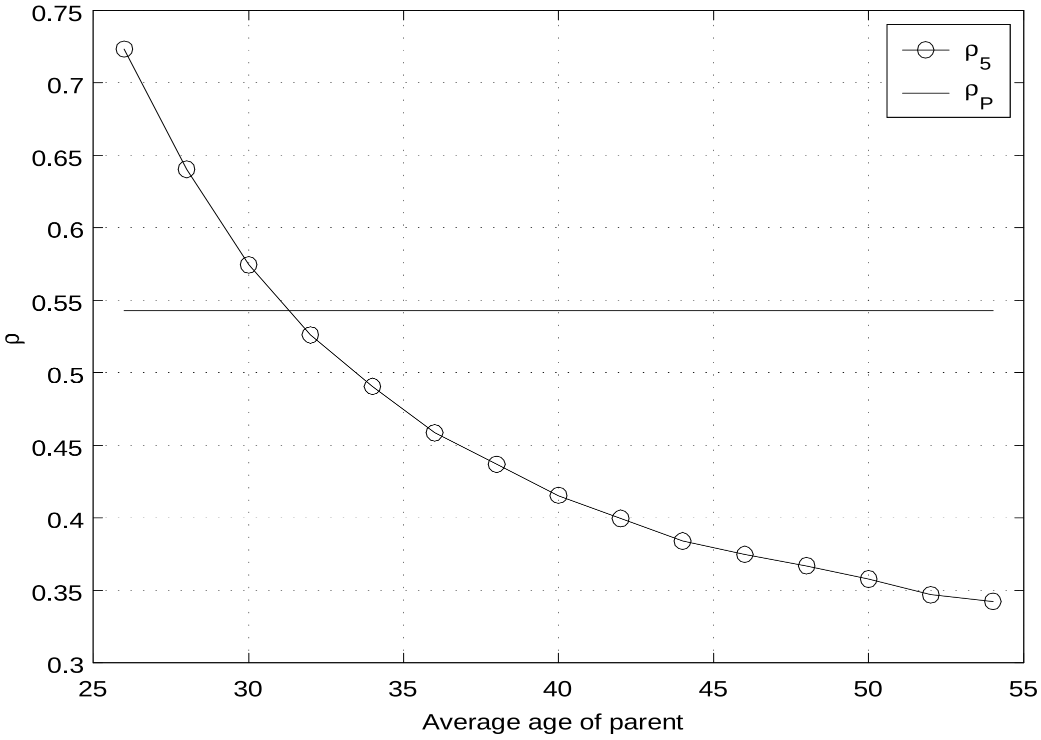

Figure 1: Parental age and intergenerational persistence

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 φ is set to the largest feasible value. The qualitative conclusions of the baseline case are confirmed. In each case, ρ_5 is below two-thirds of ρ_P. Not deleting observations with zero earnings widens the gap between ρ_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.[20] 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 is not important for estimating intergenerational persistence. It mainly governs the difference between ρ_1 and ρ_5. In what follows, I therefore set $\sigma_\varepsilon^2 = 0.063$ based on STY. The fact that the model's implied lifetime earnings persistence is robust against changes in σ_ε^2 confirms my earlier assertion that the estimation method proposed here is insensitive to serially uncorrelated measurement error.

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

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

$$Var(z(a^*)) = Var(y(a^*)) - Var(h) - \sigma_\varepsilon^2 = 0.738 \tag{5}$$

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

$$Var(z(a_1+n)) = Var(\zeta)\left[1+\alpha^2+...+\alpha^{2\,(n-1)}\right] + \alpha^{2n}\,Var(z(a_1)) \tag{6}$$

the required shock variance is given by

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

---

[20] More general earnings processes do not always imply such high persistence (Baker and Solon 2003). Exploring how this might change the findings is left for future research.

3. Find the value of φ that matches $\rho_5 = 0.37$. This is not always possible, since φ 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 φ is given by

$$\varphi^2 \leq Var(z(a_1)) \,/ Var(z(a_B)) \tag{7}$$

where $Var(z(a_B))$ is calculated from (6).

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 I consider the values $\beta = 0.9$ and 0.7 (lower values yield too small values for ρ_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, ρ_5 falls to only 0.15. However, the intergenerational persistence of lifetime earnings is 0.54, regardless of whether zero earnings observations are interpolated (ρ_F) or retained (ρ_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, φ is set to its upper bound (7) 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 (ρ_5 approximately equals two-thirds of ρ_P or ρ_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 ρ_5. The downward bias of ρ_5 is now even greater (40% or more). For the baseline DGP, the bias of ρ_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 ρ_P against ρ_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.

Table 5 also reveals that using permanent earnings (h) as a proxy for lifetime earnings yields biased estimates of intergenerational persistence. For all of the earnings processes studied here, h is substantially more persistent than lifetime earnings. The intuition is that lifetime earnings are affected by shocks that are realized at later stages of the son's life and hence not correlated with the father's earnings.

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 ρ_P and ρ_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 (2).

[INSERT TABLE 5 HERE]

[Insert figure 2 here]

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 (2), the two estimation methods are equivalent. Knowles and Hendricks estimate an individual fixed effect, $y_0(i)$, based on a regression of the form

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

over some fixed age range, where $\tau(i,a)$ 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

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

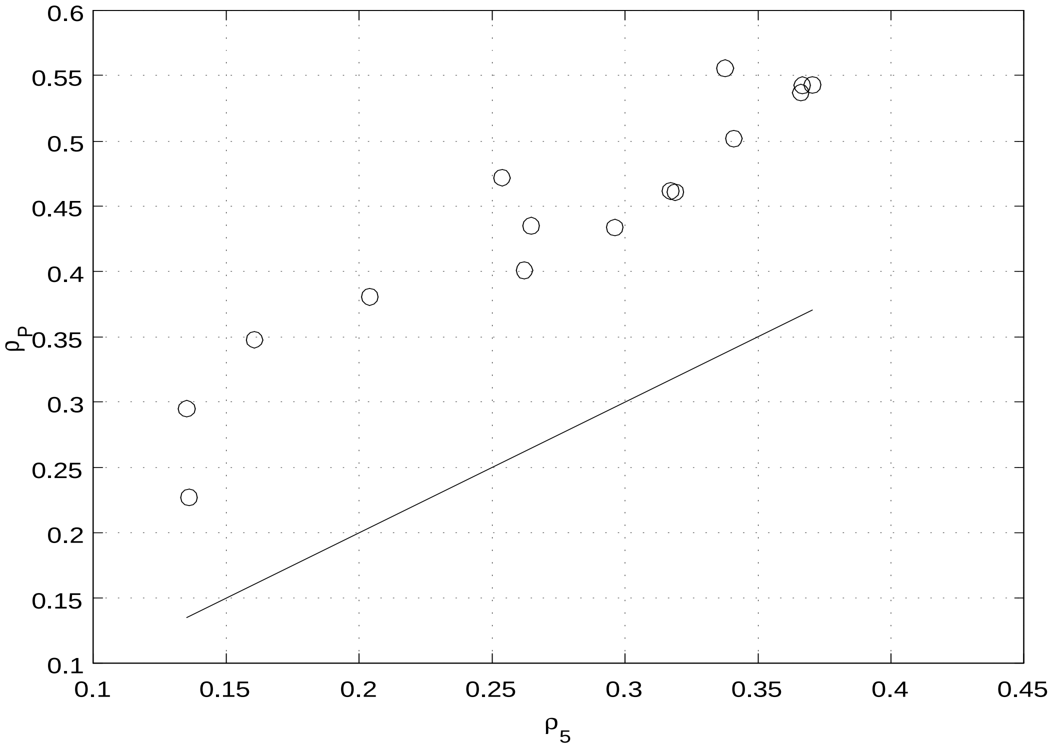

Figure 2: Sensitivity analysis

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 ρ_5. This accounts for the fact that the persistence estimates of Knowles (1999) and Hendricks (2001) are similar to estimates of ρ_5 published in the literature.[21]

5.2.2 Instrumental Variables

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

One limitation of IV estimates is that they may provide consistent estimates of the persistence of *permanent* earnings, but not of *lifetime* earnings. To illustrate this point, consider the following simple model. As an instrument for parental earnings a researcher uses a variable $w(i)$, such as education, that is correlated with individual characteristics according to

$$\begin{aligned} w(i) &= \mu\, h(i) + (1-\mu)\, z(a_{IV}, i) + \varepsilon_w \\ \varepsilon_w &\sim N(0, \sigma_w^2) \end{aligned} \tag{8}$$

The instrument may be correlated with permanent (h) and with transitory earnings (z), depending on the weight μ. Intergenerational persistence is estimated in two stages. The first stage regresses log earnings at age a_p on the instrument

$$y(a_p, i) = \xi_0 + \xi_1\, w(i) + \xi_i^*. \tag{9}$$

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 (8), μ 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 ρ_{IV} consistently estimates β. A number of authors interpret

---

[21] 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).

their IV results in this way (e.g., Solon, 1992 and Zimmerman, 1992). This might explain why IV estimates of intergenerational persistence tend to be significantly higher than OLS estimates. In the class of earnings processes studied in section 5, h must be quite persistent in order to generate the observed values of average earnings persistence (such as ρ_5 near 0.37). To the extent that IV estimates recover the persistence of h, they are upward biased as measures of lifetime earnings persistence (ρ_P or ρ_F).

The difficulty with estimating lifetime earnings persistence using IV is that ρ_c is a complicated function of all the parameters characterizing the earnings DGP. Consistently estimating ρ_c then requires an instrument that attaches the correct weights (μ) to h and z. But since these weights are unknown, little can be said about the relationship between ρ_{IV} and parameters of interest. It is not even ensured that ρ_{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 (8) 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 $z\left(a_{IV}, i\right)$ is not that large. Hence, $z(a_{IV}, i)$ is far less persistent across generations than is $z\left(a_1, i\right)$.

Of course, these numbers are mere examples. Without information on how the instruments relate to the individual characteristics that determine earnings, the relationship between ρ_{IV} and ρ_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 β (or φ). This information can then be used to help infer ρ_c using simulated earnings histories.

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 sons' 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 conditional on labor earnings being generated by a class of stochastic processes commonly used in the applied equilibrium literature. Within this class, the results are robust. However, a recent literature has proposed more general earnings models that encompass heterogeneity in mean age-earnings profiles and age dependent shock variances.

Whether the findings are robust under these models as well remains to be explored.

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

References

[1] Baker M, Solon G. Earnings dynamics and inequality among Canadian men, 1976-1992: Evidence from Longitudinal Income Tax Records. Journal of Labor Economics 2003;21(2); 289-321.

[2] Behrman JR, Taubman P. Intergenerational earnings mobility in the United States: Some estimates and a test of Becker's intergenerational endowments model. Review of Economics and Statistics 1985;67(1); 144-51.

[3] Behrman JR, Taubman P. The intergenerational correlation between children's adult earnings and their parents' income: results from the Michigan Panel Survey of Income Dynamics. *Review of Income and Wealth* 1990;36(2); 115-27.

[4] Castaneda A, Diaz-Giminez J, Rios-Rull JV. Accounting for the U.S. earnings and wealth inequality. *Journal of Political Economy* 2003;111(4); 818-57.

[5] Couch KA, Lillard DR. Sample selection rules and the intergenerational correlation of earnings. Labour Economics 1998;5(3); 313-29.

[6] Fullerton D, Rogers DL. Who bears the lifetime tax burden? Brookings: Washington, DC; 1993.

[7] Gokhale J, Kotlikoff LJ, Sefton J, Weale M. Simulating the transmission of wealth inequality via bequests. Journal of Public Economics 2001;79; 93-128.

[8] Gourinchas P, Parker J. Consumption over the life cycle. Econometrica 2002;70(1); 47-89.

[9] Grawe N. Life-cycle bias in the estimation of intergenerational earnings persistence. Mimeo. Carleton College; 2002.

[10] Haider S, Solon G. Life-cycle variation in the association between current and lifetime earnings. Mimeo. University of Michigan; 2003.

[11] Heckman JJ. Policies to foster human capital. NBER working paper #7288; 1999.

[12] Hendricks L. How do taxes affect human capital? The role of intergenerational mobility. Review of Economic Dynamics 2001;4(3); 695-735.

[13] Hubbard G, Skinner J, Zeldes S. Precautionary saving and social insurance. Journal of Political Economy 1995;103(2); 360-99.

[14] Huggett M. Wealth distribution in life-cycle economies. Journal of Monetary Economics 1996;38; 469-94.

[15] Huggett M, Ventura G. Understanding why high income households save more than low income households. Journal of Monetary Economics 2000;45(2); 361-97.

[16] Knowles J. Measuring lifetime inequality. Mimeo. University of Pennsylvania; 1999.

[17] Mazumder B. Earnings mobility in the U.S.: A new look at intergenerational mobility. Working paper #2001-18, Federal Reserve Bank of Chicago; 2001.

[18] Mazumder B. Revised estimates of intergenerational income mobility in the United States. Mimeo. Federal Reserve Bank of Chicago; 2003.

[19] Mulligan CB. Parental priorities. University of Chicago Press: Chicago; 1997.

[20] Peters HE. Patterns of intergenerational mobility of income and earnings. Review of

Economics and Statistics 1992;74(3); 456-66.

[21] Rubinstein Y, Tsiddon D. Coping with technological change: the role of ability in making inequality so persistent. Journal of Economic Growth 2004;9; 305-46.

[22] Solon G. Intergenerational mobility in the United States. American Economic Review 1992;82(3); 393-408.

[23] Solon G. Intergenerational mobility in the labor market. In: Ashenfelter O, Layard R (Eds), Handbook of labor economics, vol.3C. Elsevier: Amsterdam; 1999.

[24] Storesletten K, Telmer C, Yaron A. Consumption and risk sharing over the life cycle. Mimeo. Stockholm University; 1998.

[25] Zimmerman DJ. Regression toward mediocrity in economic stature. *American Economic Review* 1992;82(3); 409-29.