Galtonian Regression of Intergenerational Income Linkages: Biased Procedures, a New Estimator and Mean-Square Error Comparisons

Ramses H. Abul Naga*

DEEP, Université de Lausanne, CH 1015 Switzerland

June 2000

Abstract

Because the permanent incomes of parents are children are typically unobserved, the estimation of the intergenerational correlation via the use of proxy variables entails an errors-in-variables bias. By solving a system of moment equations for income observed at a given year, and a T-period average of this variable, we derive an analytical form for the signal to total variance ratio. In turn, we propose a simple estimator of the intergenerational elasticity via division of the OLS estimator by this quantity. Estimates of the intergenerational elasticity derived from a PSID sample range between 0.34 and 0.69. The averaging estimator provides intermediary values between OLS and the proposed estimator. Persistence is higher for family income measures than labor market outcomes. Estimates generally increase for moving average specifications in comparison to the assumption that measurement errors are uncorrelated. The three estimators are further examined in the light of their mean-square errors (square bias plus variance).

Keywords: intergenerational mobility, Galtonian regression, errors in variables, mean-square errors.

JEL classification: I3, J62.

---

* I am grateful to Alberto Holly, Jaya Krishnakumar and Thomas Von Ungern-Sternberg for providing helpful suggestions. I also wish to thank Richard Smith and Dirk van de gaer for discussions. None of them should be made responsible for any errors or omissions.

e-mail: Ramses.AbulNaga@hec.unil.ch

1 Introduction

For those who view inequality of incomes as being naturally inherent to the way a market economy operates, a question arises as to how individuals of different family backgrounds move about the social ladder. Do the children of poor origins, and those raised in opulence, face equal prospects of occupying various positions in the distribution of income? How many generations will it take for recent immigrants to be on average equally well off as the native population of a host country? In order to begin to address issues of this nature, one needs to formulate an empirical framework for analyzing income dynamics.

There has been a renewed interest in recent years in the estimation of the extent of income continuity across generations. The problem of estimating the intergenerational elasticity of incomes is particularly challenging since, as it stands, the variables of interest, namely the permanent incomes of parents and children, are typically unobservable. Instead, the researcher will possess a short time-series of observations on some income indicator (family income, earnings, hourly wage etc.), on the basis of which, estimation of the intergenerational elasticity is to be attempted.

Because this type of measurement error biases the ordinary least squares estimator towards zero, Behrman and Taubman (1990), Solon (1992), Zimmerman (1992), Bjorklund and Jantti (1997), Mulligan (1997), and others, have suggested to regress a measure of the child's income on the averaged income of her/his parents for the years of data available in the sample under use. The rationale underlying the method of averaging is to increase the variance ratio of permanent to observed income, and hence to reduce the asymptotic bias of the resulting estimator.

It remains nonetheless that the averaging estimator is bound to remain inconsistent in a short panel. At this stage one may attempt to model the covariance structure of the incomes of parents and children (Zimmerman, 1992; Altonji and Dunn, 1991). While such approaches may generate consistent estimators (which may also be shown to be efficient within appropriately defined classes), their validity rests on the researcher's capacity to correctly specify the moment restrictions pertaining to a system of equations. A popular alternative to such procedures consists in instrumenting parental income using family background variables such as education. This latter method is of some appeal, as it is informationally less demanding in the sense that it necessitates the choice of a single valid instrument. However, Solon (1992) has argued that instruments such as the education of the parent-head may correlate with the error term of the income transmission model, which in turn may entail inconsistent estimation of the intergenerational elasticity.

As panel data provide repeated measurements on family incomes, one may attempt to circumvent the problem of selecting valid out of equation instruments, by making use of leads and lags of parental income as within equation instruments. Thus, following Griliches and Hausman (1986), Abul Naga and Krishnakumar (1999) set out to estimate the intergenerational correlation of incomes within the context of a panel data framework with errors of measurement. Our approach in this paper is somewhat different. First, observe that the asymptotic biases of the OLS and averaging estimators are both functions of two parameters, namely the variances of permanent and transitory incomes. Next, note that provided these two variance components can be estimated, the OLS estimator may be appropriately rescaled in a way as to achieve consistency. We show in the paper that it is generally possible to obtain analytical expressions for these two variance components, as solutions to a system of two moment equations in two unknowns. Alternatively, we may observe that the OLS estimator can be viewed as a member of the family of T-period averaging estimators (the T=1 case). We may then state our approach as being an attempt to exploit the information contained in the asymptotic biases of a sequence of averaging estimators with the aim of deriving a consistent estimator of the intergenerational elasticity.

The proposed estimator in this paper is straightforward to compute and easy to interpret. Once consistent estimators are obtained for the permanent and transitory variance components, we rescale the OLS estimator through division by the estimated signal to total variance ratio in a way as to neutralize the errors-in-variables bias. We may however note at this stage that this consistency gain does not come without cost. Because the signal to total variance ratio is smaller than unity, the rescaling of the OLS estimator by this quantity will therefore increase its variance. A natural question then arises as to how to model and quantify the tradeoff between the use of an inconsistent estimator on the one hand, and a consistent statistic, with a possibly larger variance, on the other hand. One possible solution to statistical problems of this nature consists in comparing estimators in terms of their mean-square errors (squared bias plus variance), a point that that we shall discuss in more detail below.

In order to render our estimator operational we need to derive its large sample distribution. As the proposed statistic is a rescaling of the OLS estimator, its distribution is a simple linear transformation of that of the ordinary least squares estimator in an errors-in-variables environment. A study of the latter distribution, following the work of Aigner (1974), however shows that previously reported standard errors in the intergenerational mobility literature were incorrect, as they ignored a component of variance originating from the measurement error. This conclusion is likewise shown to apply to the averaging estimator. A derivation of the variance of the averaging estimator in fact reveals that an increase in T, the number of years over which parental income is averaged, has an ambiguous effect on the precision of this statistic. Thus, it may well be that while the OLS estimator has a larger asymptotic bias than the averaging estimator, its variance turns out to be smaller. Again, this observation then leaves room for OLS to dominate the averaging estimator in a mean-square error sense.

Because neither of the three estimators considered in this study can be analytically claimed to dominate any of the other two in a mean-square error sense, we offer two ways for the practitioner to think about which estimator to rely most on in empirical work. At the theoretical level, it may be noted that as the number of parent and child observations approaches infinity, the asymptotic variances of the estimators considered here vanish to zero. When working with large samples, the practitioner may therefore arguably abstract from variance considerations and rank the estimators from the least, to the most biased, in an asymptotic sense. On such grounds, the empirical analyst may wish to place most confidence on rescaled OLS (the estimator proposed in this study), and the least on unadjusted OLS.

A more cautious data analyst would however point out that, at present, most parent-child samples rarely exceed 1000 observations (see for eg. Haveman and Wolfe, 1995; table 2a). A more conservative approach to this problem would therefore consist in computing numerically the mean-square errors of the three estimators as a basis of further assessing their respective reliabilities. As sample sizes may greatly differ from one application to the next, we would in fact recommend the latter approach. We therefore provide in the paper consistent estimators for the various parameters required for the evaluation of asymptotic biases and mean-square errors. All estimators proposed here have analytical expressions, making the computation of mean-square errors a simple and straightforward exercise.

The structure of the paper is the following. Section 2, comprising four sub-sections, sets the problem of estimating the intergenerational elasticity in the context of a Galtonian model of income transmission. Section 3 presents our data, section 4 contains empirical applications, while section 5 ends the paper with some concluding comments.

In sub-section 2A we examine in some detail the consequences of measurement error. We show that the variance formulas for OLS and the averaging estimator are misspecified. We also show that consistent estimation of these quantities requires knowledge of the permanent and transitory variance components of income. In sub-section 2B we derive our estimators of the permanent and transitory variance components of income by solving a system of two moment equations in two unknowns. There, we also present our proposed estimator of the intergenerational elasticity and derive its large sample distribution. As the variance of the rescaled OLS estimator depends on the same set of parameters as in the case of OLS and the averaging estimator, our discussion in this sub-section also covers the estimation of the variances of these estimators. In the following sub-section we derive the mean-square errors of the three estimators and show why they cannot be ranked. Section 2 is closed with a discussion on how our framework may be extended in a simple way to deal with moving average-type serial correlation in the transitory component of income.

Section 4 presents an empirical application of our methodology to a US sample of parents and children extracted from the panel study of income dynamics. We have selected our observations in a way as to replicate several sampling features of data sets used in this literature (see for instance Solon, 1992 and Zimmerman, 1992). We have looked at intergenerational continuities for commonly used measures of economic status. These included the hourly wage and annual earnings of the household head, and the total family income with and without adjustment for family size. Incomes of parents were observed over the four-year period 1967-70. Incomes of children referred to the year 1991. Our estimates of the intergenerational elasticity are in the order of 0.34 to 0.69. Estimates vary according to the income definition used. Likewise, they are shown to be sensitive to the assumptions pertaining to the serial correlation in the transitory component of income. Our estimates of the signal to total variance ratio are mostly in the 0.70 to 0.84 range, suggesting that the bias of the unadjusted OLS estimator is far from being negligible.

2 Estimation and Inference

We are interested in quantifying the degree of income inheritance b in a regression of the child's permanent income h_{ic} on that of her/his parents, h_{ip}. Assuming all variables are expressed in deviations from their respective means, the Galtonian regression model is of the form

$$h_{ic} = bh_{ip} + z_i$$

where z_i is a disturbance term assumed to be uncorrelated with h_{ip}. Because the complete life movies h_{ic} and h_{ip} are typically unobserved, these variables are proxied by measurements y_{it} and x_{it} (annual earnings, family incomes, wages etc.) assumed to exhibit the classical errors in variables properties, namely:

$$y_{it} = h_{ic} + f_{it}$$
$$x_{it} = h_{ip} + e_{it}$$

The fact that h_{ic} is measured with noise does not entail biases in the estimation of b. Hence, in what follows, we will define

$$y_{it} = bh_{ip} + v_{it} \qquad (1)$$

with $v_{it} = z_i + f_{it}$, as our theoretical model.

A: Consequences of measurement errors

In contrast with the baseline model (1), the measurement model

$$y_{it} = bx_{it} + u_{it} \qquad (2a)$$
$$u_{it} = v_{it} - be_{it} \qquad (2b)$$

is subject to a specification bias in the sense that the composite error term u_{it} is correlated with x_{it} (via e_{it}).

For a given time period, the standard probability limit formula for the OLS estimator given in the literature (for eg. Solon, 1992) is

$$\text{plim}\,(\hat{b}) = bs_{pp}/(s_{pp} + s_{ee}) \qquad (3)$$

where s_{pp} is the variance of the permanent component h_{ip} and s_{ee} that of the transitory component of the parents' income.

The averaging estimator extensively used in the literature (for eg. Solon, 1992; Zimmerman, 1992; and Bjorklund and Jantti, 1997) regresses y_{it} on a time-series average $\bar{x}_i = \sum_{t=1}^{T} x_{it}/T$ on parental income, yielding an estimator $\bar{b}$ with probability limit:

$$\text{plim}\,(\bar{b}) = bs_{pp}/(s_{pp} + s_{ee}/T) \qquad (4)$$

The probability limit formula assumes that e_{it} is stationary and serially uncorrelated [1]. The appeal of the averaging estimator $\bar{b}$ can be illustrated by means of a simple numerical example. If say s_{pp}=3/4 and s_{ee} =1/4, yielding a signal to total variance ratio of 3/4, then plim($\hat{b}$)=3b/4,. With a two period average, plim($\bar{b}$)=7b/8. If T=4, plim($\bar{b}$)=12b/13 etc. As can clearly be read from (4), the bias of $\bar{b}$ will vanish as T goes to infinity. It remains though that because T is small in all data applications, constructing a consistent estimator of b may be a worthwhile task.

Because the OLS estimator (as well as $\bar{b}$) is inconsistent, its standard error is also misspecified. To see that this is so, consider its large sample distribution. Note firstly that its mean is given by the probability limit formula (3). Standard errors reported in the literature are based on

[1] See Zimmerman (1992),also Griliches and Hausman (1986).

$$\sum_{i=1}^{n}(y_{it}-\hat{b}x_{it})^2/n\sum_{i=1}^{n}x_{it}{}^2 \qquad (5)$$

where n denotes sample size. While the above formula is valid in the Gauss-Markov model, in the present context, the errors-in-variables model, it is inappropriate. Following Aigner (1974), the variance of the OLS estimator is given by

$$V(\hat{b}) = s_{**}/(s_{pp}+s_{ee}) \qquad (6)$$

where

$$s_{**} = s_{yy} - s_{xy}^{2}/(s_{pp}+s_{ee}) \qquad (7)$$

The term s_{yy} is the population variance of y, that is the variance of y evaluated in the baseline model (1):

$$s_{yy} = b^2 s_{pp} + s_{vv}$$

Likewise, s_{xy} is the covariance between y in (1) and the noisy measurement x_{it}, viz., $s_{xy} = bs_{pp}$.

Hence, gathering the various terms in (7) we obtain

$$s_{**} = s_{vv} + b^2 s_{pp} s_{ee}/(s_{pp}+s_{ee}) \qquad (8)$$

The second term on the left-hand side of (8) arises because of the measurement error problem. In absence of this, s_{ee} =0, and s_{**} reduces to the usual formula, on the basis of which, the derivation of standard errors using (5) would be correct.

Using the above results, the large sample distribution of the OLS estimator in the present context is given by

$$\sqrt{n}\left[\hat{b} - plim(\hat{b})\right] \sim N\left[0\,;\,s_{**}/(s_{pp}+s_{ee}),\right] \qquad (9)$$

For the same reasons as in the case of the OLS estimator, the estimator

$$\sum_{i=1}^{n}(y_{it}-\bar{b}\bar{x}_i)^2/n\sum_{i=1}^{n}\bar{x}_i{}^2 \qquad (10)$$

for the variance of $\bar{b}$ is misspecified. Noting that $\bar{b}$ can be treated as an estimator for a model with a regressor $\bar{x}_i$, whose variance is given by $s_{pp} + s_{ee}/T$, we may readily derive the following large sample distribution for $\bar{b}$:

$$\sqrt{n}[\bar{b} - p\lim(\bar{b})] \sim N[0; q]$$

where, q, the variance $\sqrt{n}\bar{b}$ of is given by

$$q = \frac{s_{vv}}{(s_{pp} + s_{ee}/T)} + \frac{b^2 s_{pp} s_{ee}/T}{(s_{pp} + s_{ee}/T)^2} \qquad (11)$$

The lesson to learn then, is that consistent estimators for s_{pp} and s_{ee} are required in order to identify the large sample distributions of the OLS and averaging estimators. The next sub-section of the paper deals with this task.

B: Estimation

Let x_{it} be a snapshot observation on the parents' income, and $\bar{x}_i$ be a T-period average. The variance formula for the decomposition of income into permanent and transitory components entails the following system of two equations in two unknowns:

$$V(x_{it}) = s_{pp} + s_{ee} \qquad (12a)$$

$$V(\bar{x}_i) = s_{pp} + s_{ee}/T \qquad (12b)$$

In turn, these yield as solutions

$$s_{ee} = \frac{T}{T-1}\left[V(x_{it}) - V(\bar{x}_i)\right] \qquad (13a)$$

$$s_{pp} = \frac{TV(\bar{x}_i) - V(x_{it})}{T-1} \qquad (13b)$$

Upon replacing $V(x_{it})$ and $V(\bar{x}_i)$ by their sample counterparts, consistent estimators $\hat{s}_{pp}$ and $\hat{s}_{ee}$ obtain for the variance components of x [2].

Now define l as the signal to total variance ratio:

$$l = s_{pp}/(s_{pp} + s_{ee}) \qquad (14)$$

Rewriting the probability limit formula (3) for the OLS estimator, it follows that plim($\hat{b}$)=$b\ l$. Furthermore, define $\hat{l} = \hat{s}_{pp}/(\hat{s}_{pp} + \hat{s}_{ee})$. It is a consequence of the continuous mapping theorem that provided $\hat{s}_{pp}$ and $\hat{s}_{ee}$ both converge to their population counterparts, $\hat{l}$ is also consistent for l. We propose then to estimate b by introducing a correction factor *(1/l)* to the OLS estimator in order to achieve consistency:

---

[2] Other estimators for s_{pp} and s_{ee} may be envisaged. For instance, one may replace (10b) by an equation in first differences $V(x_{it+1} - x_{it}) = 2s_{ee}$, to yield for solutions $s_{ee} = V(x_{it+1} - x_{it})/2$ and $s_{pp} = V(x_{it}) - V(x_{it+1} - x_{it})/2$. However, it is a consequence of Slutsky's theorems (Goldberger 1991, ch. 9) that the large sample distribution of b is invariant to the choice of consistent estimators of the variance components.

$$b = \hat{b} / \hat{l} \qquad (15)$$

Letting $s(x_{it})$ and $s(\bar{x}_i)$ respectively denote the sample second order moments of x_{it} and $\bar{x}_i$, this new estimator may alternatively be written as

$$b = \frac{(T-1)s(x_{it})}{Ts(\bar{x}_i) - s(x_{it})}\hat{b} \qquad (16)$$

Below we refer to this statistic as the *rescaled* OLS estimator.

Since $\hat{l}$ converges in probability to l, and $\sqrt{n}\left[\hat{b} - \mathrm{p}\lim(\hat{b})\right]$converges in distribution to (9), it is a result of Slutsky's theorem that the ratio $b = \hat{b} / \hat{l}$ has the following large sample distribution:

$$\sqrt{n}(b - b) \sim N\left[0; s_{**} / l^2 (s_{pp} + s_{ee})\right] \qquad (17)$$

We may state this result more simply by noting that since b is a simple rescaling of the OLS estimator, its large sample variance is a factor $(1/l)$ times that of $\hat{b}$.

What remains for the estimator b to become operational is the derivation of its standard error. Going back to (8), we require a consistent estimator of s_{vv}.

$$w_{it} = y_{it} - bx_{it} = y_{it} - b(h_{ip} + e_{it})$$

We have

$$\sum_{i=1}^{n} w_{it}^2 = \sum_{i=1}^{n}(y_{it} - bh_{ip})^2 + b^2\sum_{i=1}^{n} e_{it}^2 - 2\sum_{i=1}^{n}(y_{it} - bh_{ip})\ e_{it}$$

As the third term converges in probability to zero, we have that

$$p\lim \frac{1}{n}\sum_{i=1}^{n} w_{it}^{\ 2} = s_{vv} + b^2 s_{ee}$$

where, once again, the second term on the left hand side arises from the measurement error. Hence, s_{vv} may be consistently estimated via

$$\hat{s}_{vv} = \frac{1}{n}\sum_{i=1}^{n} w_{it}^{\ 2} - b^2 \hat{s}_{ee} \qquad (18)$$

where $\hat{s}_{ee}$ is the sample counterpart of s_{ee} as defined in (13a). In passing, we may note that the above estimator also provides the necessary information to

correctly estimate the standard error of the (inconsistent) OLS estimator, the latter also being a function of s_{ee}, s_{pp} and s_{vv}, as shown in equation (9).

C: Mean-square error comparisons

The estimator *b* is then the preferred one over OLS and the method of averaging in terms of consistency. The purpose of this sub-section is to further compare the three estimators in terms of mean-square errors (here defined as squared asymptotic bias plus asymptotic variance).

For a pair of unbiased estimators, mean-square error (MSE) comparisons readily translate to the familiar exercise of variance ranking. Conversely, for two estimators with identical variances, MSE contrasts amount to evaluating the estimators in terms of biases. Clearly, neither of these cases applies here, hence the need for a more in depth comparison of these estimators.

Using equations (3), (8) and (9), we may readily derive the following mean-square error formula for the OLS estimator:

$$MSE(\hat{b}) = \frac{b^2 s_{ee}^2}{(s_{pp} + s_{ee})^2} + \frac{1}{n}\left[\frac{s_{vv}}{(s_{pp} + s_{ee})} + \frac{b^2 s_{pp} s_{ee}}{(s_{pp} + s_{ee})^2}\right] \qquad (19)$$

where, in the present context, the large sample bias of $\hat{b}$ is taken as $b - p\lim(\hat{b})$ (derived on the basis of 3), and the second term of the right hand side of (19) is var($\hat{b}$) as defined in (9). For *b*, a consistent estimator, its mean-square error collapses to its large sample variance, viz. $\mathrm{var}(\hat{b}) / l^2$. On such grounds, the difference between the *MSEs* of $\hat{b}$ and b can be written as a function *D* with arguments *b*, s_{ee}, s_{pp}, s_{vv}, and *n*:

$$\Delta(b, s_{ee}, s_{pp}, s_{vv}, n) = \frac{b^2 s_{ee}^2}{(s_{pp} + s_{ee})^2} + \frac{1}{n}\left[\frac{s_{vv}}{(s_{pp} + s_{ee})} + \frac{b^2 s_{pp} s_{ee}}{(s_{pp} + s_{ee})^2}\right].\left[1 - \frac{1}{l^2}\right] \qquad (20a)$$

Or, alternatively, defining Bias($\hat{b}$)=*b*-plim($\hat{b}$),

$$\Delta = MSE(\hat{b}) - MSE(b) = \mathrm{Bias}^2(\hat{b}) + \mathrm{var}(\hat{b}).\left[1 - \frac{1}{l^2}\right] \qquad (20b)$$

The lesson to learn from either forms of (20), is that, because $0<l<1$, the first term on the right hand side is positive, while the second is negative. Hence, the price to pay in moving from $\hat{b}$ to *b*, i.e. in gaining consistency, is to have a variance increase. In general, then, the ranking of $\hat{b}$ and *b* will be ambiguous in terms of mean-square error.

However, it is possible to sign the effect of a subset of the arguments of the function *D* on the difference in the mean-square errors of the two estimators. Differentiation of D with respect to n, s_{vv}, and b is straightforward. On such grounds, the difference in mean-square errors can be shown to be increasing in n and diminishing in s_{vv}. That is, other things equal, b is to be preferred in large samples over OLS. Conversely, the larger the disturbance variance s_{vv} of the baseline model, the better $\hat{b}$ will perform vis-à-vis b in terms of mean-square error. The effect of an increase in b on the other hand cannot be signed. This is also the case for the remaining two parameters, namely s_{ee} and s_{pp}, which, together with b, appear in both the bias and variance components of (20).

Turning now to the case of the averaging estimator, define g as the variance ratio

$$g = \frac{s_{pp}}{s_{pp} + s_{ee} / T} \qquad (21)$$

Rewriting (4) using the above notation, we have $p\lim(\bar{b}) = bg$; alternatively, Bias($\bar{b}$)=b(1–g). Hence the mean-square error of $\bar{b}$ can be expressed as $MSE(\bar{b}) = b^2(1-g)^2 + q/n$, that is:

$$MSE(\bar{b}) = \frac{b^2 s_{ee}^2 / T^2}{(s_{pp} + s_{ee}/T)^2} + \frac{1}{n}\left[\frac{s_{vv}}{(s_{pp} + s_{ee}/T)} + \frac{b^2 s_{pp} s_{ee} / T}{(s_{pp} + s_{ee}/T)^2}\right] \qquad (22)$$

where q, the variance of the averaging estimator, is defined as in (11). The point to note in (22) is that, while the effect of an increase in the number of measurements T over which parental income is averaged reduces the asymptotic bias of $\bar{b}$, it nonetheless has an ambiguous effect on its variance. Hence, the effect of an increase in T is also ambiguous on the mean-square error of the averaging estimator.

This observation however entails further implications about the ranking of $\bar{b}$, $\hat{b}$ and b in terms of mean-square errors. Since OLS can be considered as an averaging estimator for which T=1, a comparison of (22) with equations (6) and (8) pertaining to the variance of the OLS estimator, shows that, in presence of finite samples, one cannot generally establish the superiority of $\bar{b}$, over $\hat{b}$ in terms of mean-square errors. While $\bar{b}$ clearly possesses a smaller asymptotic bias than $\hat{b}$ (for T>1), empirical investigations will prove useful in order to provide further guidelines concerning the relative merits of these two estimators. For the same reason that the effect of an increase in T on the variance of $\bar{b}$ cannot be signed, the MSE ranking of b and the averaging estimator cannot be established. The opposite statement would be somewhat surprising given our earlier conclusion that $\hat{b}$ and b could not generally be ranked using this latter criterion.

D: Extensions

This sub-section provides simple extensions of our framework in order to deal with estimation in presence of serially correlated transitory components of income.

The presence of a cross-section of observations on the parent family's income can be further exploited to relax the assumption that the transitory component e_{it} of x_{it} is uncorrelated over time. Previous research on the covariance structure of earnings (MaCurdy 1982, Abowd and Card, 1989, Schluter 1998) suggests that passed a certain time lag, income changes (intended to difference out time invariant components) are uncorrelated. For this reason, the parameterization of the transitory component of income by an MA(q) process may provide a useful starting point for relaxing the assumption that errors are serially uncorrelated. Other alternatives clearly exist. For instance Zimmerman (1992) adopts an AR(1) specification for this same purpose. We shall discuss this point in further detail in the final section of the paper, where we provide some directions for further research.

For now, the simplest way to examine the consequences of allowing for serially correlated errors is to note that variance formula for $\bar{x}_i$, (10), will then require modification. For example, if T=2,

$$V\left[\frac{x_{i1}+x_{i2}}{2}\right] = s_{pp} + s_{ee}/2 + \text{cov}(e_{i1}e_{i2})$$

While the estimation of covariance terms can be undertaken by making use of further sample moments, typically covariances between x_{it} and x_{is}, a simple solution here consists in taking averages for incomes observed several periods apart. If for instance we postulate that e_{it} follows an MA(1) process, equations (10) and all the subsequent developments are valid if we construct averages over $x_{it}, x_{it+2}, x_{it+4}\ldots$, etc. Likewise, for an MA(2) specification, series of the type $x_{it}, x_{it+3}, x_{it+6},\ldots$ are to be constructed.

3 Data

Our sample was extracted from the University of Michigan's Panel Study of Income Dynamics (PSID). From the first wave of the Panel (1968) we have identified families with dependent children, which we have attempted to follow up to 1992 (wave XXV). The PSID consists of two major files commonly referred to as the SRC and SEO, details of which can be found in Hill (1993). The SEO file is a sample of low income families which had participated in the Survey of Economic Opportunity in the years 1965 and 1966, and then accepted to take part in the wider survey carried out by the University of Michigan's Institute of Social Research, since 1967. The SRC component, the new sample selected by the Institute of Social Research, has been designed as a national probability sample, intended to be representative of the US population.

In the present study we have only worked with data originating from the SRC file, in an attempt to minimize the problem of homogeneity bias (Solon, 1989 and 1992) which may arise from the use of a non-random sample such as the SEO. While we recognize that studying income continuities amongst families on low income is a topic of inherent interest, we have chosen here to work exclusively with a random sample and to pursue this latter question elsewhere.

We have observed the incomes of parents over the four years period 1967-70 and those of children in 1991. In accordance with several previous US studies (eg. Solon, 1992 and Zimmerman, 1992), we have restricted ourselves to the examination of father and son linkages, leaving aside female headed households and, or, father-daughter pairs. As the labor supply decisions of men and women may be governed by different forces it is perhaps more cautious to analyze these data separately (though see Dearden, Machin and Reed, 1997 for UK evidence on the differential pattern of income inheritance across father-daughter and father-son pairs). Likewise, we have retained a single child per family in order to avoid problems of correlation across observations. We note however that this latter problem may be treated via the adoption of generalized least squares data weighting schemes.

We have looked at four commonly used indicators of economic status: total family income, family income normalized by the Orshansky needs scale, total annual earnings, and the average hourly wage of the household head. Our overall sample comprised 596 observations, though it is important to note that sample sizes vary depending on the choice of indicator being used. Table 1 summarizes our data in terms of age, hourly earnings of household heads, together with their needs-adjusted family incomes. The Consumer Price Index was used in this study in order to deflate all incomes back to 1967 dollars.

Table 1: descriptive statistics

variable	mean	standard deviation
parent head's age in 1968	40.81	9.86
child head's age in 1992	36.60	8.72
parents' income in 1967	2.60	1.55
child family's income in 1991	4.28	6.49
parent head's wage in 1967	3.85	2.21
child head's wage in 1991	4.06	3.08

Notes:
1. Incomes and hourly wages are measured in 1967 dollars.
2. Family income is normalized by the Orshansky needs scale.

Though the average ages of parents and children are fairly close (40.8 years for parents and 36.6 for children), there is a great deal of variation within each of these distributions. For this reason we have run prior regressions of log-income on the age and age squared of the household head in each given year, and we have chosen to work with the residuals from these initial regressions in order to estimate the intergenerational elasticity of income.

A further point in table 1 ought to be mentioned: the standard deviation of needs-adjusted (Orshansky) income in 1991, a 6.5 figure, is one and half times the mean of the children's distribution. This phenomenon is largely due to the presence of a family with an income in excess of 130 times its needs. The exclusion of this observation would bring the coefficient of variation down from 1.5 to 0.93, and as an approximate rule of thumb, would deflate estimates of b by 4% to 5% for the two income measures used in this study. As incomes of this size are however not unheard of in practice, we have not decided to discard this observation from our sample.

4 Applications

The purpose of the applications presented below is to shed empirical evidence on the claims made thus far in this paper. Firstly, we wish to provide new estimates of the intergenerational elasticity in the light of the rescaled OLS estimator. Next, we wish to quantify the bias of the OLS and related averaging estimator. Thirdly, we provide mean-square error comparisons between OLS, the averaging estimator and the new estimator of b, in an attempt to provide practical guidelines for future empirical work in the area. We also provide separate estimates for various concepts of economic status in order to examine the sensitivity of our results to the choice of income definition. We provide further robustness checks by estimating our models under various assumptions regarding the serial correlation of the measurement error.

Throughout our applications OLS estimators are computed from a regression of the child family's 1991 income on that of the parent family in 1967. In the first column of table 2 we report estimates of *b* using the four definitions of income considered in this study. The standard errors computed here are those derived in section 2, intended to take into account the presence of the variance component originating from the errors-in-variables problem. OLS estimates of *b* vary between 0.338 for hourly wages and 0.443 for needs-adjusted Orshansky incomes. These estimates are fairly similar to those reported by Solon (1992, table 4) whose sample is also extracted from the PSID. Solon's estimates vary between 0.294 for wages and 0.476 for Orshansky incomes (with values of 0.386 for earnings and 0.483 for total family incomes). In all cases then, these estimates are within one standard error from ours.

Table 2: estimates of the intergenerational elasticity when incomes are averaged over two years

variable	***OLS***	***AVE.***	***b***	***l***	***g***	***n***
Orshansky income	0.443 (0.057) [0.021]	0.496 (0.059) [0.006]	0.523 (0.068) [0.005]	0.843	0.915	595
tot. family income	0.425 (0.076) [0.021]	0.488 (0.080) [0.010]	0.534 (0.096) [0.009]	0.795	0.886	596
earnings	0.390 (0.057) [0.004]	0.403 (0.057) [0.003]	0.405 (0.059) [0.004]	0.962	0.981	549
wages	0.338 (0.048) [0.010]	0.360 (0.050) [0.004]	0.407 (0.058) [0.003]	0.830	0.907	549

Notes:

1. *AVE* is the method of averaging, *b* is the rescaled OLS estimator.
2. *l* is the signal to total variance ratio, *g* is the corresponding shrinkage factor for the averaging estimator. *n* denotes sample size.
3. Standard errors are reported inside curly brackets, mean-square errors are reported inside square [ ] brackets.
4. Parental income is averaged over 1967 and 1968, the child's income pertains to 1991.

The next two columns of table 2 contain estimates of *b* using the averaging estimator $\bar{b}$ and our proposed estimator *b*. In these applications, incomes of parents are averaged over 1967 and 1968 (that averaging intervenes in the computation of b can be seen via inspection of equation 16). Previous research has established that both $\hat{b}$ and $\bar{b}$ are biased toward zero, though averaging did reduce the errors-in-variables bias. Likewise, it is expected that of the three estimators, b takes on the largest numerical value as it is intended to be consistent. That b will exceed $\hat{b}$ numerically will always be the case by definition, since l, the signal to total variance ratio, is smaller than unity.

The findings of table 2 however confirm the hypothesized pattern that $\bar{b} < b$ for all four definitions of income. It may be noted that averaging over two years in the applications considered here produces an estimate of b around 10% higher when $\bar{b}$ is used instead of $\hat{b}$. The intergenerational elasticity is estimated at 0.496 for Orshansky incomes and 0.488 for total family incomes by the method of averaging, versus 0.443 and 0.425 respectively for OLS. Use of the new statistic b entails estimates of 0.523 and 0.534 respectively using normalized incomes and total family incomes. For the two concepts based on labor market outcomes, the correction introduced via calculation of the averaging estimator is less large in comparison with OLS (estimates of b increase from 0.390 to 0.403 for annual earnings of the household head, and from 0.338 to 0.360 for hourly wages). On the basis of the evidence provided by the estimator b we would however be led to conclude that the intergenerational elasticity for hourly wages is somewhat higher (a 0.407 estimate versus 0.360 using the method of averaging and 0.338 for OLS). The 0.405 estimate for earnings is however very much in line with the earlier finding obtained by the averaging estimator.

The next two columns of table 2 provide further information on the magnitudes of the biases of the OLS and averaging estimators (cf. equations 14 and 21). Estimates of l provide indications on the level of shrinkage of OLS from the population parameter b, while estimates of g serve the same purpose for $\bar{b}$. Leaving aside earnings, estimates of l are in the range of 0.80 to 0.84, while estimates of g would indicate that averaging over two years results in downwardly biased estimators, with the magnitude of the bias being in the range of 8% to 12%. For earnings, estimates of l and γ are respectively 0.96 and 0.98, which certainly stand out as being higher than in the cases of the three other indicators of income status considered in this study. We shall have more to say on this particular point as we proceed with an examination of the results of table 3.

We have argued earlier in section 2 that comparisons between biased estimators require an examination of their mean-square errors. In the context of three estimators considered in this study, it was not possible to establish analytically the superiority of either OLS, the method of averaging or b. It is therefore of interest to examine numerically the relative performance of these three estimators. In table 2 and elsewhere, we therefore report both standard errors (inside curly brackets) and mean-square errors (inside square brackets).

It may then be noted that of the three estimators $\hat{b}$ possesses generally the lowest standard error, while *b* exhibits the largest one. The tradeoff between bias reduction and variance increase, in moving away from OLS, appears to be empirically warranted in the sense that both $\bar{b}$ and *b* possess significantly lower mean-square errors. The exception to the rule once again arises when we examine earnings linkages. There, all three estimators produce very similar solutions and, as a result, their mean-square errors are within the same order of magnitude.

In table 3 we replicate the estimations of table 2 taking four-year averages of parental income. Once again, our benchmark estimator is that of an OLS regression of the child's 1991 income on that of his parents in 1967. For this reason, the first column of table 3 replicates the same OLS estimates of *b* as the corresponding column of table 2. However, because averaging is now undertaken over a four-year horizon, (see equations 13) estimates of s_{pp} and s_{ee} may change, hence also altering the standard error and MSE of $\hat{b}$. Estimates of the intergenerational elasticity resulting from the averaging estimator over a four-year horizon are all higher than the corresponding estimates for *T*=2. The estimate of the intergenerational elasticity rises from 0.496 to 0.527 for needs-adjusted incomes, from 0.488 to 0.513 for total family incomes, from 0.403 to 0.408 for earnings and from 0.360 to 0.379 for wages. With the exception of earnings, where the increment is negligible, these findings confirm earlier results by Behrman and Taubman (1990), Solon (1992), Zimmerman (1992) and others, that averaging over a longer time horizon reduces the bias of the estimator $\bar{b}$.

Table 3: estimates of the intergenerational elasticity when incomes are averaged over four years

variable	*OLS*	*AVE.*	*b*	*l*	*g*
Orshansky income	0.443 (0.057) [0.017]	0.527 (0.061) [0.005]	0.559 (0.072) [0.005]	0.792	0.939
tot. family income	0.425 (0.076) [0.027]	0.513 (0.084) [0.009]	0.569 (0.102) [0.010]	0.746	0.922
earnings	0.390 (0.057) [0.003]	0.408 (0.057) [0.003]	0.400 (0.058) [0.003]	0.974	0.993
wages	0.338 (0.048) [0.011]	0.379 (0.052) [0.004]	0.433 (0.061) [0.004]	0.781	0.935

Notes:

1 AVE is the method of averaging, *b* is the rescaled OLS estimator.

2 l is the signal to total variance ratio, g is the corresponding shrinkage factor for the averaging estimator.
3 Standard errors are reported inside curly brackets, mean-square errors are reported inside square brackets.
4 Parental income is averaged over the years 1967-70, the child's income pertains to 1991.

With the exception of the case of earnings, we may note that estimates of the intergenerational elasticity provided by the statistic b are also generally higher. For Orshansky incomes the estimate of b rises from 0.523 to 0.559; for total family incomes the estimate increases from 0.534 to 0.569, while for hourly wages the increment is from 0.407 to 0.433. In all three cases, the increase amounts to one half of a standard error.

Another way of stating that b is generally higher is to note that estimates of l (with the exception of earnings) are also somewhat smaller that those of table 2. Figures for the variance ratio of permanent to observed income are in the range of 0.75 to 0.79, implying that unadjusted OLS estimates would require an upward correction of approximately 25 to 33%, depending on the choice of income definition adopted. Likewise, estimates of g based on four-year averages would require multiplication by a factor of 1.065 to 1.085.

Standard errors for $\bar{b}$ are somewhat higher when averaging is undertaken over four years instead of two years. However, our calculations imply that the variance increase is offset by a reduction in (square) bias, in the sense that, in moving from T=2 to T=4, there is a small decrease in mean-square error. The mean-square error of b is simply equal to its variance, a decreasing function of l. Because estimates of l in table 3 are somewhat smaller than those of table 2, this in turn results in a slight mean-square error increase for b. However, in choosing between the three estimators, MSE calculations once again tilt the balance against $\hat{b}$. Despite exhibiting a smaller variance, OLS is estimated to possess 2.5 to 3 times the mean-square error of the other two estimators depending on the definition of income status considered.

It is in the case of earnings that none of these conclusions appear to hold. The signal to total variance ratio l is estimated at 0.962 in table 2, and at 0.974 in table 3. Taken at face value, these results would imply that the bias of the OLS estimator is very small when examining earnings continuities. We have two reasons for calling this conclusion to doubt. Firstly, it may be noted that for the three other indicators the range of estimates is substantially lower, and (combining the findings of tables 2 and 3) in the order of 0.75 to 0.84. Furthermore, other available estimates in the literature do not offer evidence of l being so close to unity. Bowles (1972, table A1) reports estimates ranging between 0.70 and 0.83 (for various income concepts), while Zimmerman (1992, table 14) estimates this ratio to be 0.73 for wages and 0.66 for earnings.

Though we do not possess a full explanation for the rather high estimate of l in the case of earnings, we have found that it is sensitive to the inclusion of individuals who supply few hours of labor annually. For instance, excluding 16 observations for parent and child heads who supply less than 500 hours results in an estimate of 0.915 for l, with $\hat{b}$ =0.412(0.059), $\bar{b}$ =0.463(0.061) and *b*=0.451(0.065). Likewise, if we exclude individuals who supply fewer than 1000 hours of work annually, leaving us with a sample of 500 parent and child pairs, l falls further to 0.851, with the three estimators taking the following values $\hat{b}$ =0.414 (0.058), $\bar{b}$ =0.450 (0.061) and *b*=0.487 (0.068). In this latter case the estimate of l for earnings falls closer in line with the other related estimates of table 3.

Table 4: estimates of the intergenerational elasticity with serially correlated measurement errors

	MA (1)			MA (2)		
	AVE.	***b***	***l***	***AVE.***	***b***	***l***
Orshansky income	0.509 (0.061) [0.014]	0.620 (0.080) [0.006]	0.714	0.514 (0.061) [0.015]	0.624 (0.080) [0.006]	0.709
tot. family income	0.502 (0.083) [0.021]	0.628 (0.112) [0.013]	0.676	0.492 (0.084) [0.034]	0.689 (0.123) [0.015]	0.616
earnings	0.392 (0.057) [0.003]	0.399 (0.058) [0.003]	0.978	0.422 (0.058) [0.004]	0.422 (0.061) [0.004]	0.924
wages	0.364 (0.051) [0.007]	0.454 (0.064) [0.004]	0.745	0.378 (0.051) [0.009]	0.474 (0.067) [0.005]	0.713

Notes:

1. *AVE* is the method of averaging, *b* is the rescaled OLS estimator and l is the signal to total variance ratio.

2 Standard errors are reported inside curly brackets, mean-square errors are reported inside square [ ] brackets.

3 Parental income is averaged over 1967 and 1969 for the MA(1) model, and over 1967 and 1970 for the MA(2) specification.

In table 4 we relax the assumption that the transitory component of parental income is uncorrelated over time, by adopting a moving average specification for e_{it}. We have estimated an MA(1) process (by averaging incomes over 1967 and 1969) as well as an MA(2) specification (which averages incomes over 1967 and 1970). Higher order moving average processes may further be examined by constructing longer time-series of observations on parental income, however in the present study we have limited our time span to four years of measurement.

We may now note that b estimates the intergenerational elasticity to be above 0.6 in the context of Orshansky and total family incomes. The estimates for the two specifications are broadly similar in the case of needs-adjusted incomes, whereas for total family incomes b increases from 0.628 to 0.689 (approximately one half of a standard error) in moving from the MA(1) to the MA(2) model. We may also note that these estimates are higher than those of tables 2 and 3, which were obtained under the assumption that transitory incomes were uncorrelated over time. It may also be noted that in both the MA(1) and MA(2) specifications b produces a higher estimate in the case of wages, though the increment is less large than for Orshansky and total family incomes.

The case of earnings stands again apart with the MA(1) model producing for b very similar solutions to those of tables 2 and 3. We may observe nonetheless that in this case the MA(2) model departs slightly from the other results, where b is estimated at 0.422 instead of 0.400.

For the averaging estimator, we may note that estimates of the intergenerational elasticity are only marginally higher for the MA(2) model in comparison to the MA(1) specification (the estimate for family incomes is in fact smaller than in the MA(1) specification). As averaging in table 4 is undertaken over two periods, a natural comparison for $\bar{b}$ with the benchmark assumption that errors are uncorrelated, ought to be performed by cross-examining estimates of table 2 (also based on two-year averages). It may be noted here that differences in estimates are only minor in comparison to the contrasts depicted by the rescaled OLS estimator b.

We may further discriminate between $\bar{b}$ and b in the case of serially correlated measurement errors by examining the mean-square errors reported in table 4 for these two estimators. Unlike the earlier results of tables 2 and 3, mean-square error contrasts between these two estimators now go in favor of b despite its higher dispersion in comparison to $\bar{b}$. This is of course another way of observing that differences between these two estimators tend to be larger once we abandon the assumption that errors are uncorrelated over time.

In an attempt to summarize the lessons learned from these empirical applications, we plot in figures 1 and 2 the range of estimates of the intergenerational elasticity produced by $\bar{b}$ and b. One immediate lesson is that the choice of income definition does matter. From both figures 1 and 2 we may tentatively conclude that measures based on total family resources will depict a higher elasticity estimate than those based on labor market outcomes. Bearing in mind that our earnings estimates must be handled with caution, we may note that the 0.45 line in figure 1, and the 0.50 line in figure 2, separate the range of estimates based on family incomes from those derived from labor market outcomes. By comparing figures 1 and 2, we may also note that for a given income concept the range of estimates tends to be tighter when looking at the averaging estimator (with the exception of earnings where the brackets are of the same size).

Finally, in figure 3 we plot the range of estimates of the signal to total variance ratio l. These plots provide an overall view of the extent of the bias of unadjusted OLS estimators of the intergenerational elasticity. The amount of shrinkage of an OLS estimator may be quite large, in the 20% to 40% bracket when looking at family incomes, and 16% to 29% in the case of Orshansky incomes and wages. Our estimates also suggest that the bias of the OLS estimator is fairly small when examining earnings continuities, but we have also shown that estimates of l are fairly sensitive to the inclusion of individuals who supply a small amount of hours annually on the labor market.

5 Conclusions

The large sample biases of both OLS and the method of averaging are functions of the same two parameters, namely the permanent and transitory variance components of income. By solving a system of two moment equations for the variance of income at a given year, and a T-period average of this same variable, we have derived separate analytical expressions for the permanent and transitory variance components. In turn, we have proposed a simple consistent estimator of the intergenerational elasticity of income via division of the OLS estimator by the estimated signal to total variance ratio. We have also provided some straightforward extensions of our framework that cover the case of moving average type serial correlation in the errors of measurement.

Previously reported standard errors for OLS and the related averaging estimator ignored a component of variance originating from the error of measurement, and were accordingly misspecified. The paper has also provided appropriate variance formulations for these two estimators, based on a study of the distribution of the OLS estimator in an errors-in-variables context. A derivation of the variance of the averaging estimator has in fact shown that the number of years over which parental income is averaged has an ambiguous effect on the dispersion of this statistic. In our empirical applications, we have in fact found it to be generally the case that OLS exhibits a lower standard error than the averaging estimator. Thus, while the averaging estimator is unambiguously preferred over OLS on grounds of its smaller asymptotic bias, this conclusion does not carry over in the context of variance rankings.

Because the signal to total variance ratio is smaller than unity, the consistent estimator we have proposed in this paper by definition exhibits a larger variance than the OLS estimator. As a means of formalizing the existing tradeoff between bias reduction and variance increase, we were led to compare the three estimators we have examined in this study in terms of their mean-square errors. It was shown that no general ranking is available between any given pair of estimators. However, in increasingly larger samples, the variances of the three estimators vanish to zero. On such grounds, mean-square error rankings collapse to a simpler exercise of comparing estimators in terms of their asymptotic biases. In large samples then, a case can be made for preferring the rescaled OLS estimator over the method of averaging, and the latter over the unadjusted OLS estimator.

In practice, however, parent and child samples rarely exceed 1000 observations. For this reason, we have cautioned against the above large sample reasoning, and have suggested as an alternative guideline to compute mean-square errors numerically. The purpose of our empirical applications was to provide new estimates of the intergenerational elasticity, to evaluate the biases of the OLS and averaging estimators, and to compute mean-square errors for the three estimators within the context of a medium size US sample.

Our estimates of the intergenerational elasticity range between 0.34 and 0.69. Underlying this variation are factors related to the estimation method, the income definition, and the assumptions underlying the serial correlation in the transitory component of income. Estimates range between 0.34 and 0.44 for OLS, between 0.36 and 0.53 for the method of averaging, and between 0.40 and 0.69 for the rescaled OLS estimator. As a general rule, measures based on total family resources tend to depict more persistence than those based on labor market outcomes. For instance, with reference to the rescaled OLS estimator, the range of estimates pertaining to total family and Orshansky incomes are both above 0.5, while those derived from annual earnings and hourly wages are below this value. The 0.45 figure separates estimates based on family resources and those derived from labor market measures when the averaging estimator is employed instead.

We have also found that a relaxation of the assumption that errors of measurement were serially uncorrelated tended to result in higher estimates of the intergenerational elasticity. This pattern is more pronounced in the case of MA(2) models over their MA(1) counterparts, in comparison to the benchmark specification that transitory income is serially uncorrelated. Top of the range estimates for all four income concepts (0.62 for Orshansky incomes, 0.69 for total family incomes, 0.42 for annual earnings, and 0.47 for hourly wages) were in fact obtained from the rescaled OLS estimator under the MA(2) specification.

Differences between OLS estimates and those provided by our proposed estimator, rescaled OLS, may be accounted for by the size of the variance ratio of permanent to total income. Estimates of the signal to total variance ratio vary between 0.71 and 0.84 for needs adjusted family incomes, between 0.62 and 0.80 for total family incomes, between 0.92 and 0.98 for earnings, and between 0.71 and 0.83 for hourly wages. These results imply that, depending on the choice of income definition, the asymptotic bias of the OLS estimator may be quite large. Our results suggest that the OLS estimator shrinks the population elasticity by 20% to 40% in the case of total family incomes, and by approximately 16% to 29% when working with hourly wages and Orshansky incomes. It is in the case of earnings continuities that the bias of the OLS estimator would appear to be small. However, we have also noted that earnings based estimates of the signal to total variance ratio are fairly sensitive to the sampling of individuals who work a small number of hours annually.

Our numerical mean-square error calculations were undertaken in order to provide a common ground for comparing estimators with differing biases and variances. While OLS will often exhibit the smallest variance, the magnitude of its bias is such as to render it the least preferable of the three estimators on grounds of mean-square error. When transitory income is taken to be uncorrelated over time, the averaging and rescaled OLS estimators perform broadly alike according to the mean-square error criterion. The rescaled OLS estimator depicts a larger dispersion in the data, while the averaging estimator is biased towards zero. Our calculations suggest that in a sample of 550 to 600 observations the variance increase of the rescaled OLS estimator makes up for the bias reduction in comparison to the averaging estimator.

This latter conclusion however no longer holds when we relax the assumption that the transitory component of income is uncorrelated over time. In our MA(1) and MA(2) model estimates, the averaging estimator possesses 1.7 to 2.5 times the mean-square error of the rescaled OLS estimator. Setting aside the case of annual earnings where all estimators provide very similar solutions, the gains from consistent estimation become apparent in our data when errors of measurement are taken to be moving average processes.

At a more general level, we may note that estimation problems similar in nature to those discussed in this paper are bound to occur in micro-models where permanent income features as an explanatory variable. A related literature on siblings correlations in earnings seeks to quantify the importance of family and community background variables in the determination of economic attainment. By examining the parallel estimation problems underlying this literature and the ones on income transmission (and these are clearly spelled out in Solon, 1999), we note that our framework may equally be made suitable for providing a new perspective on the analysis of siblings correlations in economic outcomes.

There are some certainly more complex error structures which we have not covered in our discussion. Typically, the transitory component of income may follow other laws of motion than the MA specification considered here. Auto-regressive, or a mixture of auto-regressive and moving average (ARMA) specifications would necessitate a reformulation of the moment equations on the basis of which the signal to total variance ratio is to be derived. Rethinking estimation in the light of ARMA error processes is certainly of great value, as it nests within its framework both the moving average and auto-regressive specifications considered thus far. Estimation of these more general error processes would however require longer time series on parental income than the ones considered to-date in the empirical literature on intergenerational mobility. Nonetheless, we believe this could well be a fruitful area for further research.

References

Abowd J. and D. Card (1989): “On the Covariance Structure of Earnings and Hour Changes”, *Econometrica* 57, 411-445.

Abul Naga R. and J. Krishnakumar (1999): “Panel Data Estimation of the Intergenerational Correlation of Incomes”, Discussion Paper 9910, DEEP, University of Lausanne.

Aigner D. (1974): "MSE Dominance of Least Squares with Errors of Observation", *Journal of Econometrics* 2, reprinted in Aigner D. and A. Goldberger eds (1977): *Latent Variables in Socio-Economic Models*, Amsterdam, North Holland.

Altonji J. and T. Dunn (1991): “Relationships Among Family Incomes and Labor Market Outcomes of Relatives”, *Research in Labor Economics* 12, 269-310.

Atkinson A., J. Maynard and C. Trinder (1983): *Parents and Children: Income in Two Generations,* London, Heineman.

Behrman J. and P. Taubman (1990):"The Intergenerational Correlation Between Children's Adult Earnings and their Parents' Income: Results from The Michigan Panel of Income Dynamics", *Review of Income and Wealth* 36, 115-27.

Bjorklund A. and M. Jantti (1997): "Intergenerational Mobility in Sweden Compared to the United States", *American Economic Review* 87, 1009-1018.

Bowles S. (1972): "Schooling and Inequality from Generation to Generation", *Journal of Political Economy* 80, S219-S251.

Dearden L., S. Machin, and H. Reed (1997): “Intergenerational Mobility in Britain”, *Economic Journal* 107,47-66.

Haveman R. and B. Wolfe (1995): “The Determinants of Children’s Attainments: A Review of Methods and Findings”, *Journal of Economic Literature* 33, 1829-1878.

Goldberger A. (1991): *A Course in Econometrics*, Boston MA, Harvard University Press.

Griliches Z. (1986): "Economic Data Issues", in Griliches Z. and M. Intriligator eds: *A Handbook of Econometrics* vol. III, Amsterdam, North Holland.

Griliches Z. and J. Hausman (1986): "Errors in Variables in Panel Data", *Journal of Econometrics* 31, 93-118.

Hill M. (1993): *The Panel Study of Income Dynamics: A User’s Guide,* New York, Sage Publications.

MaCurdy T. (1982): “The Use of Time Series Processes to Model the Error Structure of Earnings in Longitudinal Data Analysis”, *Journal of Econometrics* 18, 83-114.

Mulligan C. (1997): *Parental Priorities and Economic Inequality*, Chicago, Chicago University Press.

Schluter C. (1998): “Income Dynamics in Germany, the USA, and the UK, Evidence form Panel Data” CASE paper no. 8, STICERD, London School of Economics.

Solon G. (1989): “Biases in the Estimation of the Intergenerational Earnings Correlation”, *Review of Economics and Statistics* 71, 172-174.

Solon G. (1992): "Intergenerational Income Mobility in the United States", *American Economic Review* 82, 393-408.
Solon G. (1999): “Intergenerational Mobility in the Labor Market”, in Ashenfelter O and D. Card eds: *Handbook of Labor Economics*, volume 3, 1761-1800, Amsterdam, Elsevier.
Zimmerman D. (1992): "Regression Toward Mediocrity in Economic Stature", *American Economic Review* 82, 409-429.

Figure 1: Range of estimates of *b* derived from the averaging estimator

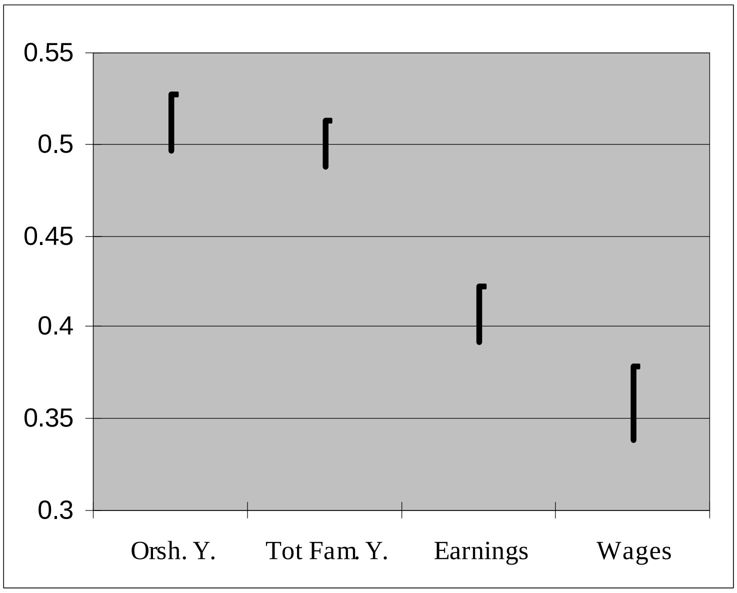

Figure 2: Range of estimates of *b* derived from the modified OLS estimator

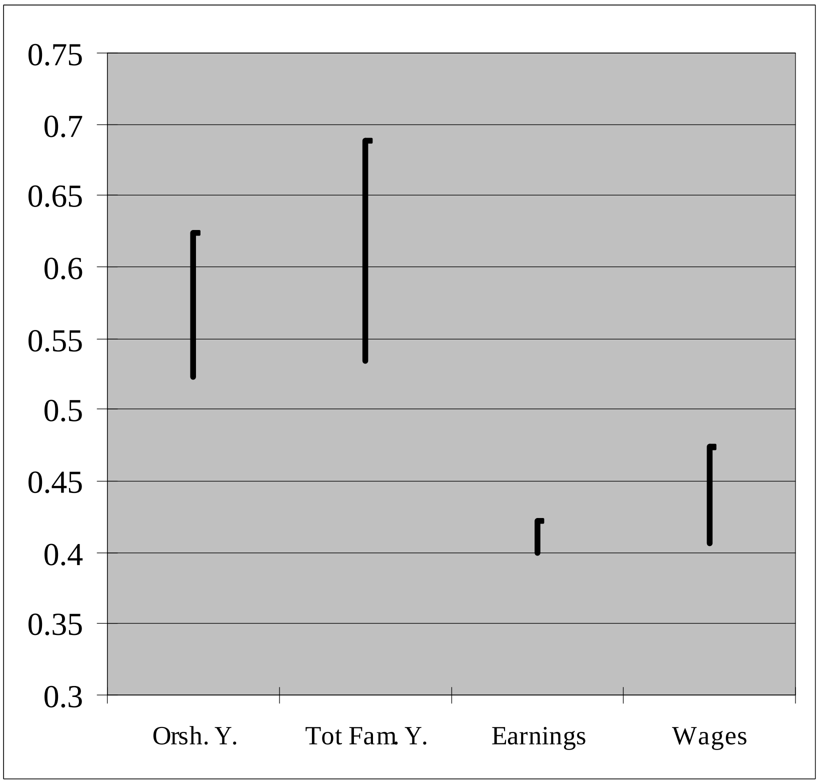

Figure 3: Range of estimates of l, the signal to total variance ratio