Estimating the Intergenerational Correlation of Incomes: An Errors in Variables Framework

Ramses H. Abul Naga[*]
DEEP, Université de Lausanne, CH-1015 Lausanne, Switzerland.

June 15, 2000

Revised, June 2000

Abstract

Because the permanent incomes of parents and children are typically unobservable, the estimation of the intergenerational correlation of incomes is usually carried out via averaging methods or instrumentation. In this paper we take the permanent income of the parent family to be unobserved, but we assume that a model for its determinants is known to the researcher. In turn, this leads us to propose two related estimators for the intergenerational correlation: a two-stage least squares procedure and a more e¢cient MIMIC estimator. The MIMIC framework also provides estimates for the determinants of permanent income and the variance parameters required to evaluate the bias of the OLS estimator. Using a US sample of parents and children we provide estimates for the intergenerational correlation ranging between 0.30 and 0.78. The bias of the OLS estimator is estimated to be in the order of 40%.

Keywords: Intergenerational mobility, errors in variables.
JEL codes: I3, J62.

[*] I wish to thank Tony Atkinson, Robin Burgess and Jean Imbs for discussions. I am grateful to Alberto Holly, Dirk Van de gaer, Thomas Von Ungern-Sternberg and an anonymous referee for detailed comments on the paper. This research was supported by a grant no. 12-46751-96 from the Swiss National Science Foundation. Any errors are my sole responsibility.

1 Introduction

One question which has been high on the agenda of researchers and policy makers alike during the past three decades concerns the extent to which children inherit the income position of their parents. Some fifteen years ago, Becker and Tomes (1986) reviewing evidence from studies of several advanced economies, concluded that the intergenerational correlation of incomes was most probably in the order of 0.20. Such a conclusion implies that the process of regression to the mean is indeed rapid: the grandchild of an individual with twice the average income today would only be expected to be 4% above the mean income level two generations from now. This conclusion has however been challenged by Behrman and Taubman (1990), Solon (1992), Zimmerman (1992) and Björklund and Jantti (1997) amongst others, who derived estimates of the intergenerational correlation in the United States ranging between 0.30 and 0.80. With this upper bound estimate of 0.80, regression to the mean would be considerably smaller than what Becker and Tomes had suggested, and returning to our earlier example, the grandchild of an individual with double the average income would be 64% above the mean in his generation, rather than 4% above.

Having a reliable estimate of the intergenerational correlation may clearly alter the way we think about justice and the extent to which a society comes close to an objective of equal opportunities. The problem of estimating the intergenerational correlation 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, an hourly wage etc.), on the basis of which estimation of the intergenerational correlation is to be attempted. One important consequence that follows from the use of an error-ridden explanatory variable is that the corresponding regression coefficient is biased towards zero (Griliches, 1986). As a result, earlier attempts to quantify the intergenerational correlation which failed to address the measurement error problem (such as the studies surveyed by Becker and Tomes) systematically under-estimated the extent of income inheritance. In the more recent studies, it was suggested to average incomes over several years in a way as to limit the effect of biases resulting from measurement error. Solon (1992) and Zimmerman (1992) have also opted for an instrumental variables estimation strategy, with variables such as education being frequently chosen to instrument the error-ridden measurement on the income of parents.

While instrumental variables estimators are now routinely being used in this literature, a point still remaining largely undiscussed is how information from various instruments may be combined to formulate more efficient estimators of the intergenerational correlation than those considered thus far, and what further information regarding the process of income transmission may be drawn from such an exercise, beyond the estimation of the correlation parameter.

In order to address this question, in the present paper we consider a situation whereby the permanent income of the parent family is unobserved, but we assume that the researcher can specify a model for its determinants. This scenario leads us to propose two related estimators for the intergenerational correlation. The first of these is a two-stage least squares procedure consisting in (1) predicting parental permanent income using the model for its determinants, and (2) regressing the child's income measurement on the predicted permanent income of the parent family. The estimator resulting from this two-stage procedure is analytically equivalent to instrumenting the error-ridden measurement on the permanent income of the parent family using the full set of determinants of this variable. Thus, formulating a model for the determinants of permanent income is useful in that it describes the full set of instruments that may be used in the estimation of the intergenerational correlation.

Beyond this however, joint estimation of a model of the determinants of parental permanent income and a Galtonian regression of the child's income on that of the parents allows for efficiency gains over the standard two-stage least squares/instrumental variables estimator, but also provides estimates for other key parameters in the analysis of intergenerational income linkages. The resulting system of equations has the structure of a model with Multiple Indicators and Multiple Causes on an unobserved variable; the so-called MIMIC model of Jöreskog and Goldberger (1975). Joint estimation of this system of equations entails cross-equation restrictions between the slope coefficients of the reduced form and also covariance restrictions between the disturbance terms. Information from these two types of restrictions is combined in the MIMIC framework in order to produce a more efficient estimator of the intergenerational correlation than that provided by the standard two-stage least squares/instrumental variables estimator. As a by-product of the MIMIC framework, we obtain estimates for the model of the determinants of permanent income. Also, using a standard analysis of variance decomposition, we obtain a direct measure of the bias of the ordinary least squares estimator of the intergenerational correlation–the so called signal to total variance ratio.

The plan of the paper is the following. In section 2 we present a standard family resource allocation problem and the resulting model of income transmission. The discussion there provides the starting ground for the econometric methods developed in section 3, and also sets an agenda for future research. In section 4 we summarize the main features of our US data extracted from the University of Michigan's Panel Study of Income Dynamics. In section 5 we present our empirical results. We provide a range of estimates for the intergenerational correlation of family incomes varying between 0.30 and 0.78, and we estimate the bias of the ordinary least squares estimator to be in the order of 40%. We also provide related estimates for the intergenerational correlation in the earnings of fathers and sons. In section 6 we conclude the paper with a summary of the main points and directions for further research.

2 An economic framework

The empirical literature on intergenerational mobility has provided a wide range of estimates on the elasticity of a child's income with respect to that of her/his parents. One natural question to ask is how to relate the estimates of the intergenerational correlation to the structural parameters of a behavioural model of income transmission. Likewise, examining an economic model of parental decisions may provide clues into the choice of estimation strategy, the selection of instruments and some likely biases derived from the speci…cation of linear functional forms. We take up these points here. The next section builds on the present discussion by considering further complications related to the measurement of permanent income.

Becker and Tomes (1986) consider a model where parents maximize a utility function $'[c_p; \mu(¿(1 + r) + I_c)]$ de…ned over their consumption c_p and their child's welfare $\mu(,)$. The arguments of the child's welfare function are $¿$, the size of the transfer (inter-vivo gifts plus bequests), and I_c, the child's permanent earnings. The parameter r denotes the interest rate [1]. Parents

[1] Because earnings and the transfer enter additively into the child's utility function, the above model is often referred to as the Wealth model. An alternative formulation, the Separable Earnings Transfers (SET) model proposed by Behrman et al. (1982), takes the

allocate their resources m between their consumption, an investment h_c in the child's human capital, and the transfer ¿. Mulligan (2000) further introduces uncertainty regarding earnings and the returns to …nancial assets in the child's generation. The same luck parameter l_c is assumed to govern earnings outcomes and the returns to assets, and is introduced as a multiplicative of ¿ and I_c. Hence, the family allocation problem may be formalized as follows:

$$\max \ ' \ f c_p; E[\mu(\text{¿}(1 + r) + I_c) \exp(l_c)]g \tag{1}$$

$$s{:}t{:} \quad c_p + h_c + \text{¿} = m \tag{2a}$$

$$I_c = I(h_c; A_c) \tag{2b}$$

$$\text{¿} \ \text{¸} \ 0 \tag{2c}$$

where $E(:)$ denotes the expectations operator. The constraint (2b) is a human capital production function which links permanent earnings I_c to the level of human capital investment h_c and the child's innate ability A_c. The function I is increasing in its two arguments, and is further assumed to have a decreasing …rst derivative with respect to h_c. The constraint (2c) states that the transfer given to the child cannot be negative.

Consider …rst the case where (2c) is not binding. This is the situation where parents' resources are large enough for them to invest in the child's human capital at the margin where the expected rates of return on education and …nancial assets are equal. This e¢ciency result [2] likewise obtains if credit markets are assumed to be perfect, in which case parents maximize the objective function (1) subject to the constraints (2a) and (2b) only.

form $'[c_p; \mu_1(I_c); \mu_2(\text{¿}(1 + r))]$. As discussed by Behrman (1997), the Wealth and SET models produce di¤erent conclusions in multiple children families regarding how parents' choices act to reinforce or reduce innate di¤erences in the earnings capacities of their o¤-springs.

[2] The e¢ciency property of the educational investment in fact requires some further assumptions such as the existance of no public goods or externalities at the level of the household. See Behrman et al. (1995) for a discussion.

A fairly large family of functional forms pertaining to the family utility function (1) and the human capital production technology (2b) produce explicit solutions for the above problem (see Becker and Tomes, 1986; Behrman, 1997 and Mulligan, 2000). For instance consider the following simple speci…cations:

$$' = \log(c_p) + aE\,f\log[\mu(¿(1+r) + I_c)\exp(l_c)]g \qquad (1')$$

$$I(h_c; A_c) = h_c^g \exp(A_c) \qquad (2b')$$

where $a > 0$ is the degree of parental altruism and $0 < g < 1$ is the elasticity of labour income with respect to human capital. Then, the e¢cient human capital investment is such that

$$h_c^{g\,¡\,1} = \frac{1+r}{g\exp(A_c)} \qquad (3)$$

In particular, using the earnings function $I_c = h_c^g \exp(A_c)$, we obtain a semi-logarithmic relation between permanent income and ability: $\log(I_c) = \text{const} + ±A_c + l_c$ where const is a constant term, and $± = 1=(1\ ¡\ g)$. Assume without loss of generality that the ability and luck variables have zero means, and de…ne $´_c = \log(I_c)\ ¡\ E[\log(I_c)]$. That is, $´_c$ is the logarithm of the child's permanent earnings measured in deviation from its mean. We may then write in more convenient notation the following relation:

$$´_c = ±A_c + l_c \qquad (4a)$$

Assuming the child's parents were also raised in wealthy families, we may obtain a comparable permanent earnings-ability relation

$$´_p = ±A_p + l_p \qquad (4b)$$

In accordance with the terminology used in the empirical literature on intergenerational mobility, we shall more simply refer to $´_p$ and $´_c$ as the permanent incomes of the parent and child families.

Without the introduction of further assumptions, equations (4a) and (4b) would entail a zero correlation between the permanent incomes of parents and children. Becker and Tomes (1986) assume ability follows an auto-regressive process of the type

$$A_c = \beta A_p + l_A \qquad (5)$$

where l_A is a zero-mean disturbance taken to be uncorrelated with A_p. Together, equations (4a), (4b) and (5) may be used to derive the following population relation between the permanent incomes of unconstrained parent and child families:

$$\eta_c = \beta \eta_p + \delta l_A + l_c - \beta l_p \qquad (6)$$

Han and Mulligan (1998) draw to our attention the fact that

$$E(\eta_p \eta_c)/E(\eta_p^2) = \frac{\delta^2 \beta \, var(A_p)}{\delta^2 \beta \, var(A_p) + var(l_p)} < \beta \qquad (7)$$

where $var(x)$ denotes the variance of the variable x. That is, because the relation (6) does not obey the classical regression properties, in that $E[\delta l_A + l_c - \beta l_p | \eta_p] \neq 0$, a hypothetical regression of the child's permanent income on that of her parents' would underestimate β. Estimation of (6) by the method of instrumental variables would necessitate the selection of instruments which do not correlate with either of the luck terms l_p and l_c, or the uninherited component l_A of the child's ability. Using the relation (3), we may note that the child's human capital may not be used to instrument η_p, since h_c is a function of A_c, and hence of l_A. However, if a similar relation to (3) is assumed to hold for the human capital endowment h_p of the parent family, h_p may serve as an instrument for η_p. Further requirements will be stipulated as we examine in the next section biases related to the measurement of permanent income.

The above account of intergenerational income transmission changes considerably when the credit constraint (2c) is binding. Then, parents with limited resources will invest a smaller amount in the child's education than the efficient level calls for. As in this case it is generally not possible to obtain an explicit relation between the child's education and the parents' permanent income, equation (6) providing the family link between η_p and η_c takes the following form for wealth constrained families:

$$\eta_c = f(\eta_p; I_p; A_c) + A_c + I_c \tag{8}$$

A number of practical problems occur in empirical work if one were to take into account the fact that a proportion of the population may be wealth constrained. Note firstly that because the relation (8) is implicit, the use of an arbitrary functional form to estimate the intergenerational correlation may result in more or less severe specification biases. The estimation of (8) by more flexible techniques such as kernel-based methods is not problem free since the composite error term $A_c + I_c$ is likely to be correlated with $f(.)$ (via the association between A_c and η_p).

Han and Mulligan (1998) point out that the task of estimating the intergenerational correlation β could be greatly simplified if one could ascertain which families are wealth constrained. If this were the case, one could in principle use the complementary sample (families which invested in their child's education at an efficient scale) to run a linear regression (6) in order to estimate β. However, because in practice the data analyst does not observe whether a family is wealth constrained, Han and Mulligan caution against the use of ad-hoc methods in splitting the data between the two regimes, as these may entail sample selection biases.

We may in fact note that the presence of wealth constrained families complicates the analysis further. The implicit function (8) is likely to take on a different form depending on whether grand-parents themselves were wealth constrained when deciding how much to invest in the education of parents. If little is known in practice about whether a child's parents had a binding resource constraint, even less information is likely to be available about the economic situation of grand-parents.

Assuming the constraint (2c) is not binding for grand-parents, Han and Mulligan provide simulations on the effect of estimating the linear model (6) when a proportion of parents are not able to undertake efficient investments in their children's education. Though their findings depend on the assignment of numerical values to an extended set of structural parameters, a general conclusion is that the specification of the linear relation (6) results in an underestimate of β in presence of some wealth constrained families[3].

In our conclusions section we mention some potential strategies for estimating β in presence of wealth constrained families. We differ this discussion to the last section of the paper as there exist further estimation problems related to the measurement of permanent income, to which we presently turn.

3 Methods

In this section we examine the estimation of the linear model of income transmission (6) in the light of further biases related to the measurement of permanent income. The present model differs from the Galtonian model considered earlier in the literature in the sense that, in the population relation, the composite error term is correlated with the regressor, i.e. parental permanent income. As a result, when snapshot observations on earnings, family incomes, etc., are used to proxy permanent incomes, probability limits for OLS and related estimators all take on different forms.

We thus begin this section by deriving new probability limit formulae for OLS, the method of averaging, and the instrumental variables estimator. We then move on to consider the estimation of the intergenerational correlation under the assumption that a model of the determinants of permanent income is known to the researcher, though η_p is unobserved. This, in turn, leads us to propose two related estimators for β: a two-stage least squares procedure and a more efficient MIMIC-type estimator. The methods proposed here are intended to produce consistent estimators of the intergenerational correlation under the assumption that families face no quantity constraints in the credit market for human capital loans.

Suppose then we were interested in estimating the intergenerational correlation of incomes, but, instead of observing the permanent incomes η_{pi} and η_{ci}, all we observed were snapshot measurements of the type

$$x_{it} = \eta_{pi} + u_{it} \tag{9}$$

$$y_{it} = \eta_{ci} + \varepsilon_{it} \tag{10}$$

[3] A key parameter in isolating the effect of resource constraints mentioned by Han and Mulligan is the elasticity of substitution of consumption.

where $i = 1; \ldots; n$ indexes individuals and $t = 1; \ldots; T$ indexes the time period. The error terms u_{it} and ϕ_{it} have zero means and are assumed to be identically independently distributed, and uncorrelated with their related permanent income components. Below we shall refer to u_{it} and ϕ_{it} as transitory incomes. In contrast with the population regression (6), the measurement model takes the form[4]:

$$y_{it} = \beta x_{it} + \delta l_{Ai} + l_{ci} - \beta l_{pi} - \beta u_{it} + \phi_{it} \tag{11}$$

The component $\delta l_{Ai} + l_{ci} - \beta l_{pi}$ of the disturbance term can be thought of as a time-invariant family effect, while the remaining two terms, $-\beta u_{it}$ and ϕ_{it} vary over time. In what follows we shall make the assumption that all random variables are stationary and homoscedastic, so that $var(x)$ will denote the population variance of x_{it}, and $cov(xy)$ will denote the population covariance between x_{it} and y_{it}, etc. Assuming all five components of the composite regression disturbance are mutually uncorrelated, the probability limit of the OLS estimator of the measurement model has the form

$$plim(\hat{\beta}_{OLS}) = \beta - \frac{\beta[var(l_p) + var(u)]}{var(x)} \tag{12a}$$

or, using (4b) and (9) to obtain the relation $x_{it} = \delta A_{pi} + l_{pi} + u_{it}$, we may write the alternative form

$$plim(\hat{\beta}_{OLS}) = \beta - \frac{\beta[var(l_p) + var(u)]}{\delta^2 var(A_p) + var(l_p) + var(u)} \tag{12b}$$

The OLS estimator is biased towards zero because the measurement x_{it} is correlated with two components of the error term: l_{pi} and u_{it}. In the population model (6) transitory variations in income are inexistent–the term $var(u)$ vanishes, so that we obtain the result of Han and Mulligan; our equation (7). In earlier treatments of the Galtonian model, the luck terms l_{pi} and l_{ci} were not present. Setting $var(l_p) = 0$ in (12a) reduces the expression to the probability limit formula given in Solon (1992) and others.

[4] By analogy with the definition of η_p and η_c in the previous section, we assume variables x, y, and z (below) all have zero means.

A related estimator, the method of averaging, consists in regressing the child's income measurement y_{it} on a time series average $\bar{x}_i = \sum_{t=1}^{T} x_{it}/T$ on the income of parents (Behrman and Taubman, 1990; Altonji and Dunn, 1991; and others). In the present context, the probability limit of the estimator $\hat{\beta}_t$ resulting from such a procedure is given by [5]:

$$\text{plim}(\hat{\beta}_t) = \beta - \frac{\beta[\text{var}(l_p) + \text{var}(u)/T]}{\delta^2 \text{var}(A_p) + \text{var}(l_p) + \text{var}(u)/T} \qquad (13)$$

Since the ratio on the right hand side of (12b) is smaller than unity, it follows that the asymptotic bias of $\hat{\beta}_t$ is smaller than that of the ordinary least squares estimator. However, if the variance of the measurement error u_{it} is small in comparison to that of the luck term l_{pi}, the correction obtained by the method of averaging may be fairly limited.

A variant of the method of averaging, used by Behrman and Taubman (1990) and Mulligan (2000) regresses $\bar{y}_i = \sum_{t=1}^{T} y_{it}/T$ on $\bar{x}_i$. That is, both the dependent and explanatory variable become time-series averages. Let $\hat{\beta}_{ave}$ denote the estimator resulting from this regression. For the purpose of interpreting our empirical findings it will be useful to derive an algebraic relation between $\hat{\beta}_{ave}$ and $\hat{\beta}_t$. Define $a_i = \bar{x}_i / \sum_i \bar{x}_i^2$. We may then write

$$\hat{\beta}_t = \sum_i \bar{x}_i y_{it} / \sum_i \bar{x}_i^2 = \sum_i a_i y_{it} \qquad (14)$$

and

$$\hat{\beta}_{ave} = \sum_i \bar{x}_i \bar{y}_i / \sum_i \bar{x}_i^2 = \sum_i a_i \bar{y}_i \qquad (15)$$

i.e.

$$\hat{\beta}_{ave} = \sum_i a_i (y_{i1} + \ldots + y_{iT})/T = \sum_{t=1}^{T} \hat{\beta}_t / T \qquad (16)$$

That is, $\hat{\beta}_{ave}$ is the arithmetic mean of the various estimates $\hat{\beta}_t$, $t = 1, \ldots, T$. Its probability limit will therefore be identical to that of $\hat{\beta}_t$, as given in equation (13). However, under the assumption that the transitory income

[5] All estimators derived in this paper take the sample size going to infinity with T fixed.

components u_{it} and $Á_{it}$ are stationary, $\hat{^{}}_{ave}$ will be a more e¢cient estimator than $\hat{^{}}_t$.

Below we discuss how auxiliary information may be used in order to identify the intergenerational correlation of incomes. Before we close the discussion on the use of OLS related inference methods, we mention a further result formalized by Klepper and Leamer (1984). By observing that the regression of x_{it} on y_{it} in the measurement model (11) is also subject to an errors-in-variables problem, we may use the inverse of the OLS estimator of the reverse regression to obtain an upper bound estimate of ¯. Together with the OLS estimate of the regression of y_{it} on x_{it}, these two …gures span the range of values any consistent estimator of ¯ can take.

Suppose then that together with the measurements y_{it} and x_{it}, the data analyst possesses auxiliary information z_i on a correlate of parental permanent income. It is assumed that z_i is uncorrelated with the three error terms pertaining to the child's data in the composite disturbance of the measurement model (11). Thus setting $E(z_i l_{ci}) = E(z_i l_{Ai}) = E(z_i Á_{it}) = 0$, the instrumental variables estimator would have a probability limit of the form

$$\text{plim}(\hat{^{}}_{IV}) = ¯ \; ¡ \; ¯ \; \frac{\text{cov}(zl_p) + \text{cov}(zu)}{\text{cov}(zx)} \tag{17}$$

In this general form the instrumental variables estimator could be upwardly or downwardly inconsistent depending on the sign of the ratio inside the square brackets. Three cases of the above formula may be considered. The …rst of these is a situation where z_i does meet the orthogonality requirements $E(z_i l_{pi}) = E(z_i u_{it}) = 0$, so as to render the instrumental variables estimator consistent:

$$\text{plim}[\hat{^{}}_{IV} \; j \; \text{cov}(zl_p) = \text{cov}(zu) = 0] = ¯ \tag{18a}$$

Next, assume the instrument z_i is uncorrelated with the luck term l_{pi}, but allow it to correlate with transitory income u_{it}. For instance, z_i could be an unanticipated family transfer. Then, noting that $E(z_i x_{it}) = ±E(z_i A_{pi}) + E(z_i l_{pi}) + E(z_i u_{it})$, we would have

$$\text{plim}[\hat{^{}}_{IV} \; j \; \text{cov}(zl_p) = 0] = ¯ \; ¡ \; ¯ \; \frac{\text{cov}(zu)}{±\text{cov}(zA_p) + \text{cov}(zu)} \tag{18b}$$

In this second case, it would not be possible to sign the large sample bias of the IV estimator since, depending on the de…nition of z, $\text{cov}(zA_p)$ and cov(zu) could be of different signs. Finally, let $E(z_i u_{it}) = 0$, but allow z_i and the luck term to correlate. We then have

$$\text{plim}[\hat{\beta}_{IV} | \text{cov}(zl_p) = 0] = \beta - \beta \left[\frac{\text{cov}(zl_p)}{\delta \text{cov}(zA_p) + \text{cov}(zl_p)} \right] \qquad (18c)$$

It is instructive to consider the much discussed case in the literature where z_i is the education of the parent family head. If educational attainment may be taken as a reasonable proxy for human capital endowments h_{pi} of section 2, then the education of the parent family head will be a valid instrument provided this variable does not correlate with either of the luck term l_{pi} and transitory income u_{it}. On theoretical grounds it may not be immediate to think of correlations between human capital and transitory income. The framework of the previous section would also be indicative of a zero correlation between h_{pi} and l_{pi}. However, if for a moment we allow educational attainment to be a function of the luck term l_{pi}, it may be reasonable to assume in this case that $\text{cov}(zA_p)$ and $\text{cov}(zl_p)$ are both positive. The use of education as an instrument would then produce a downwardly inconsistent estimator. This result is to be contrasted with the conclusion in Solon (1992) that education could produce an upwardly inconsistent estimator of β. Solon's result is based on a situation where the education of the parent family head features as an explanatory variable in the theoretical model, but not in the measurement model. Accordingly, when it is used to instrument the income proxy in the measurement model, it treats the errors-in-variables problem, but the IV estimator is inconsistent because of an omitted variables bias. Solon however states that once η_{pi} features as an explanatory variable, it is quite likely that the residual effect of education is small, rendering the bias of the instrumental variables estimator negligible. Likewise, if we maintain an assumption of zero correlation between educational attainment and the stochastic terms l_{pi} and u_{it}, education also produces a consistent instrumental variables estimator in our framework [6].

More generally, consider a situation whereby permanent income is not observed, but a model for the determinants of η_{pi} is known to the researcher. As permanent income is a function of ability (equations 4), we may, in a

[6] More specifically assume S_i denotes educational attainment, such as $S_i = h_{pi} + \xi_i$, where ξ_i is a disturbance term. Then for S_i to be a valid instrument, both components h_{pi} and the measurement error ξ_i must be uncorrelated with transitory income and the luck term l_{pi}.

first stage, write our model in terms of observable correlates of A_{pi}. Let $Z_i = [z_{i1}, \ldots, z_{ik}]'$ be a set of observed correlates of ability, and let υ_i denote a disturbance which is orthogonal to Z_i such that

$$A_{pi} = \alpha' Z_i + \upsilon_i \tag{19}$$

Using (4b), our model of the determinants of η_{pi} would take the form

$$\eta_{pi} = \delta\alpha' Z_i + \delta\upsilon_i + l_{pi} \tag{20}$$

i.e.,

$$\eta_{pi} = \gamma' Z_i + \varepsilon_i + l_{pi} \tag{21}$$

where $\gamma' = \delta\alpha'$ and $\varepsilon_i = \delta\upsilon_i$. By substituting (21) into (9), we obtain

$$x_{it} = \gamma' Z_i + \varepsilon_i + l_{pi} + u_{it} \tag{22}$$

while upon replacing (10) into the population model (6), we obtain for the child

$$y_{it} = \beta\eta_{pi} + \delta l_{Ai} + l_{ci} - \beta l_{pi} + \phi_{it} \tag{23}$$

As η_{pi} is unobserved in the above equation, we propose first to derive a predictor η^*_{pi} for the parent family's permanent income, and then to regress y_{it} on η^*_{pi}. In particular, let $\eta^*_{pi} = E(\eta_{pi}|Z_i)$. Define Z as the $n \times k$ matrix with ith line Z_i' and X_t as an $n \times 1$ vector with ith element x_{it}. Then, our predictor may be constructed in the sample via the linear projection of X_t on Z:

$$\eta^*_{pi} = Z_i'(Z'Z)^{-1}Z'X_t \tag{24}$$

Next, we regress y_{it} on η^*_{pi} (by ordinary least squares) to obtain the following two-stage least squares estimator for β:

$$\hat{\beta}_{TS} = \sum_i \eta^*_{pi} y_{it} / \sum_i \eta^{*2}_{pi} \tag{25}$$

It would be useful to provide an interpretation for the suggested two-stage estimation procedure in order to make the link with other estimators used in the paper and elsewhere in the literature. Writing $P_Z = Z(Z'Z)^{-1}Z'$ as the projection matrix with properties $P_Z' = P_Z$ and $P_Z P_Z = P_Z$, $\hat{\beta}_{TS}$ may equivalently be written as

$$\hat{\beta}_{TS} = (X_t' P_Z X_t)^{-1} X_t' P_Z Y_t \tag{26}$$

where Y_t is the $n \times 1$ vector with ith element y_{it}. Thus, $\hat{\beta}_{TS}$, the OLS estimator of the regression of y_{it} on η^*_{pi}, amounts to instrumenting x_{it}, in the regression of y_{it} on x_{it} using the full set of instruments contained in Z_i. This method then is one way of interpreting previous uses of IV procedures in the literature where variables such as education in Solon (1992), an index of socio-economic status in Zimmerman (1992), or a combination of both in Dearden et al. (1997), have been used in order to instrument the proxy variable for parental permanent income.

As the two-stage least squares estimator using k instruments $z_{i1}, \ldots, z_{ik}$ can be expressed as a weighted sum of the k IV estimators using each z_{ij}, it follows from (17) that for $\hat{\beta}_{TS}$ to be consistent, each of the k instruments must be orthogonal to both the luck term l_{pi} and transitory income u_{it}. As this may appear as being a fairly demanding requirement, in our empirical applications we successively test the orthogonality requirement for each new instrument by means of Sargan tests of overidentifying restrictions (Godfrey, 1988 pp. 167-176).

A related estimator proposed by Dearden et al. (1997) consists in (i) predicting both the permanent incomes of the child and parent families, and (ii) regressing η^*_{ci} on η^*_{pi}. Though their population model differs from our equation (6), we may summarize the essence of the "prediction approach" along the following lines. By analogy with Z_i, define W_i as a set of instruments for the child's permanent income. Likewise, define W as the $n \times k$ matrix with ith line W_i' and $P_W = W(W'W)^{-1}W'$. Then, as in (24), we may obtain

$$\eta^*_{ci} = W_i'(W'W)^{-1}W'Y_t \tag{27}$$

Let $\hat{\beta}_{PA}$ denote the OLS estimator of the regression of η^*_{ci} on η^*_{pi} (the prediction approach). We have:

$$\hat{\beta}_{PA} = \sum_i \eta^*_{pi}\eta^*_{ci} / \sum_i \eta^{*2}_{pi} = (X_t' P_Z X_t)^{-1} X_t' P_Z P_W Y_t \tag{28}$$

The rationale underlying the prediction approach is however unclear to us. Observe that errors-in-variables biases only occur as a result of mismeasurement of η_{pi}; the fact that η_{ci} is measured with noise does not entail problems of this sort. Instrumenting η^*_{ci} does not seem necessary to us, and if anything, complicates further the requirements to establish the consistency of $\hat{\beta}_{PA}$ in comparison to the conditions pertaining to the two-stage least squares estimator.

There is however a case worth mentioning where the prediction approach collapses to $\hat{\beta}_{TS}$. Consider a situation whereby the researcher uses the same variables Z_i to predict the child's permanent income (i.e. the case $W_i = Z_i$). We would then have $\eta^*_{ci} = Z_i'(Z'Z)^{-1}Z'Y_t$ and expression (24) for η^*_{pi}. Since $P_Z' = P_Z$ and $P_Z P_Z = P_Z$, the estimator $\hat{\beta}_{PA}$ resulting from setting $P_W = P_Z$ in (28) would produce the two-stage least squares estimator.

Finally, we consider the estimation of β by a more efficient variant of the instrumental variables methodology. Rewrite equation (22) as (29a) below, and substitute (21) into (23) to obtain (29b):

$$x_{it} = \gamma' Z_i + \varepsilon_i + l_{pi} + u_{it} \tag{29a}$$

$$y_{it} = \beta\gamma' Z_i + \beta\varepsilon_i + l_{ci} + \delta l_{Ai} + \phi_{it} \tag{29b}$$

Equation (29a) is an empirical model for the determinants of permanent income, while (29b) is a Galtonian regression where η_{pi} has been replaced by its determinants. Joint estimation of the reduced form (29a) and (29b) allows for efficiency gains over the estimator $\hat{\beta}_{TS}$ for two reasons. Firstly, note that in the above reduced form $E(x_{it}y_{it}|Z_i) = \beta var(\varepsilon)$. That is, by pooling the information contained in the covariance matrix of the residuals with that contained in the $2 \times k$ matrix $\begin{pmatrix} \gamma' \\ \beta\gamma' \end{pmatrix}$ of slope coefficients, we may arrive at a more efficient estimator of β. The other cross-equation restriction which is exploited in the estimation of the above system is the fact that in the reduced forms (29a) and (29b) the coefficients on Z variables ought to be multiples of one another [7].

It should be noted that (29) has the structure of a system of equations with multiple indicators and multiple causes on a latent variable; a MIMIC model in the terminology of Jöreskog and Goldberger (1975)[8]. In our applications section we shall therefore refer to the estimator of β derived from

[7] In more technical terms, this is equivalent to the statement that the $2 \times k$ matrix of reduced form coefficients ought to satisfy a unit rank condition.

[8] The multiple indicators being x_{it} and y_{it}, the multiple causes Z_i, and the latent variable η_{pi}.

(29) as the MIMIC estimator. The above estimator is implemented using a LISREL routine under the assumption of multivariate normality, details of which can be found in Jöreskog and Goldberger (1975).

There have been several attempts in the literature to quantify the bias of the ordinary least squares estimator of β. Bowles (1972) draws upon evidence from studies prior to his work in order to evaluate the variances of the permanent and transitory components of income so as to construct a corrected estimate of the intergenerational correlation. Zimmerman (1992) specifies a covariance model for the various measurements x_{it} and y_{it} from which he derives (within sample) estimates of the permanent and transitory variance components of income.

Going back to our probability limit formula (12a) for the OLS estimator, we may write $\text{plim}(\hat{\beta}_{OLS}) = \beta\lambda$ with

$$\lambda = 1 - \frac{\text{var}(l_p) + \text{var}(u)}{\text{var}(x)} \tag{30}$$

By analogy with earlier uses of the terminology, we shall refer to λ as the signal to total variance ratio. The MIMIC system (29) provides estimates of β, γ, the variance of ε_i, and variances for the composite error terms $l_{pi} + u_{it}$ and $l_{ci} + \delta l_{Ai} + \phi_{it}$. Hence, as a by-product of the MIMIC framework, we may also obtain an estimate of the bias of the ordinary least squares estimator.

4 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. The SRC, 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 this study we have only worked with data originating from the SRC file, in an attempt to minimize the problem of homogeneity bias (Solon, 1989) which may arise from the use of a non-random sample such as the SEO[9]. We have observed the incomes of parents over the five year period 1967-71 and those of children twenty years after, between 1987 and 1991. 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, as developed for instance in Abul Naga and Krishnakumar (1999).

We have looked at intergenerational continuities in total family incomes and in the annual earnings of the household head. It is frequent in this literature to restrict sampling to fathers and sons only, and to study for the major part continuities in earnings (for example see Solon, 1992). However, when examining a broad concept such as total family income, it makes somewhat less sense to exclude from the sample female headed parent families and, or, daughters. Accordingly, we have retained parent and child pairs from both sexes, giving us a sample of 592 observations. From these, we have also worked with a reduced sample of 369 observations on fathers and sons, for which we have examined continuities in earnings.

TABLE 1 ABOUT HERE

Table 1 summarizes our 592 data points in terms of age and total family income of parents and children, together with other socio-economic characteristics of parent families. Incomes of parents pertain to 1967, and those of children to 1991. The Consumer Price Index was used in this study to deflate all incomes back to 1967. Though the average ages of parents and children are fairly close (43.7 years for parents and 39.7 years 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 age and age squared of the household head in each given year, and we have chosen to work with residuals from these initial regressions in order to estimate the intergenerational correlation of incomes.

[9] Furthermore, the SEO does not sample families with resources in excess of twice their needs (as defined by the US based Orshansky needs scale). Therefore, wealth constraints are likely to be more present in the SEO data. Thus, we have found it wiser to leave the analysis of SEO data for further research which develops appropriate techniques to estimate models similar to (8). See our conclusions section for further discussion.

Because early career earnings are highly noisy measures of long run income status, we did not sample parent and child family heads under the age of 25. Missing data on earnings arise when individuals do not participate in the labour force (typically because of retirement, unemployment or disability). Cases of missing data on total family income are however more rare, as this variable includes receipts from earnings, transfers, and other types of income. The advantage of working with family income over earnings is that it allows us to obtain a wider coverage of the population, especially in the case of the unemployed who are expected to be in the bottom deciles of the distribution of income. One problem, on the other hand, with the use of total family income has to do with the fact that some children are married while others are single. Some single children may expect to marry in the near future (and perhaps experience some income gains), while some married children may subsequently divorce (and may possibly undergo some losses of income). On such grounds, the use of family incomes adds further noise in the regressions in comparison to the earnings of the household head[10].

5 Results

In this section we present estimates of the intergenerational correlation of family incomes using (a) OLS and the method of averaging, (b) the two-stage least squares procedure and (c) the MIMIC estimator. We also examine somewhat more briefly intergenerational linkages in the earnings of fathers and sons using our smaller sample of observations for individuals reporting non-zero labour incomes.

A: OLS and the method of averaging

In table 2 we report estimates of β using OLS and the method of averaging. The line entry defines the period over which the child's income is measured, while the column entry pertains to the income of parents. Accordingly, the cells 1987/1967, 1987/1971, 1991/1967 and 1991/1971 are all OLS regressions with incomes measured over single year periods. The columns 1967-8 to 1967-71 respectively pertain to two to five-year averages of parental income. In the last line of the table, the entry 1987-91 signifies that the child's

[10] A solution to this problem may consist in using needs-adjusted family incomes. As family composition changes with the years, the use of needs-adjusted family incomes may introduce noise in both the numerator and denominator of the measurement, and may accordingly exacerbate existing errors-in-variables biases.

income, the dependent variable, is averaged over the …ve-year period starting in 1987 and ending in 1991.

Looking at the …rst line of the table, where the dependent variable is the child's 1987 income, we may note that the two separate OLS estimates 1987/1967 and 1987/1971 are approximately 0.30 with standard errors of 0.05. Averaging however does change the picture. A two-period average on parental income increases the estimate to 0.34, a three-year average estimates ¯ at 0.36, while four and …ve-year averages produce estimates in the order of 0.38. Solon (1992, table 3) provides similar evidence on the e¤ect of averaging earnings of parents. His single year estimates for his balanced sample range between 0.29 and 0.37 and his …ve-year average produces a 0.41 …gure.

TABLE 2 ABOUT HERE

In the next line of table 2 we repeat the earlier exercise, with the di¤erence that the child's income is now measured …ve years later, in 1991. It may be noted that, as in the study of Reville (1995), estimates of ¯ are substantially revised upwards as children age. The simple OLS estimates 1991/1967 and 1991/1971 are respectively 0.43 and 0.38 in contrast with the 0.30 …gure for the regressions of 1987. A two-year average on parental income estimates ¯ at 0.47, and a …ve-year average produces an estimate of 0.50. The corresponding …gures in the …rst line of the table are 0.34 and 0.38 respectively. The fact that children are …ve years older in 1991 eliminates some noise in their incomes due to search and matching problems in the labour market. In young samples, children with low levels of skills may appear to do well in comparison to those who undertake long years of education and training. However, in later years di¤erences between the incomes of skilled and unskilled individuals become more apparent, and the relative positions of children begin to look more like those of their parents[11].

In the last line of the table we average the children's incomes over the …ve-year period starting in 1987. The resulting estimates provide middle range …gures between the corresponding …ndings of the 1987 and 1991 regressions. They also exhibit lower standard errors than the earlier estimates.

[11] Reviewing the study of Reville (1995), Solon (1999) comments: "If, among sons in their twenties, the ones destined for higher long-run earnings are about to experience more rapid earnings growth than the ones destined for lower long-run earnings, the measurement error in the earlier years is 'mean-reverting' and causes a downward inconsistency in the estimation of the intergenerational elasticity."

Recall from (16) that $\hat{\beta}_{ave}$ is the arithmetic mean of estimators $\hat{\beta}_t$ with $t = 1987, \ldots, 1991$ in the present context of our applications. Thus, β is estimated at 0.37 when the income of parents is measured in 1967, and 0.43 when the latter variable is averaged over the period 1967-1971. Averaging the income of parents thus does alleviate biases resulting from errors in the measurement of parental permanent incomes but, as a general rule, the overall picture that emerges from the results is that the intensity of income inheritance becomes more pronounced as children become older.

We have mentioned in our methods section that by running reverse regressions it may be possible to obtain upper bound estimates for the coefficient β. For instance, if we regress the parents' 1967-1971 income average on the corresponding 1987-1991 average for the child, the slope coefficient is estimated at 0.29 (with a standard error of 0.03). Using the Klepper and Leamer result discussed in section 3, it follows that β lies in the interval [0.43 ; 3.50] (where the upper bound is the inverse of 0.29). As this upper bound is of little informational content about β, it is necessary to examine estimates resulting from other estimation strategies for this parameter.

B: Two-stage least squares procedure

As discussed in section 3, the two-stage least squares procedure is formally equivalent to instrumenting the measurement on parental income using the variables Z_i. In table 3 we consider nine instruments: the education bracket of the parent family head (the variable edu) and eight dummy variables. The variable unskill takes a unit value if the parent family head is an unskilled worker (and zero otherwise). The dummy south pertains to families living in the southern part of the United States, house takes a unit value if the family owns its dwelling and medins signifies that the parent family has a medical insurance scheme (covering all household members). The dummy ill is set equal to unity when the family head suffers from some physical or nervous illness, nonwhite indexes nonwhite family heads and smoke signifies that the parent head is a smoker. Finally, the union dummy takes a unit value when the family head is a member of a workers' union. In the regressions of table 3 the dependent variable is defined as the child's total family income in 1991, the explanatory variable is the total family income of parents in 1967 (the year prior to 1968 when the survey was started) and the instruments Z_i pertain to 1968.

TABLE 3 ABOUT HERE

We may note from the results of table 3 that IV estimates of ¯ range between 0.67 and 0.78. In comparison to the corresponding OLS estimate (the cell 1991/1967 of table 2), these …gures imply a substantial correction. One important issue in assessing the validity of these results is the extent to which the instruments used in the estimations meet the required orthogonality condition with respect to the regression error term. Taking the dummy variable for unskilled labour as our benchmark instrument, we have estimated ¯ at 0.67. Taking education as an overidentifying instrument, the estimate of the intergenerational correlation rises to 0.71. The Sargan test for assessing the validity of education as an instrument is a Â^2 variate with a single degree of freedom. The test statistic takes a value of 0.165 (last column of the table), while the critical value of the test for a 5% probability of type-I error is 3.84. Accordingly, the test does not reject the assumption that education is a valid instrument.

In results not shown, we have tested individually each of the remaining seven instruments. Of these, only the union dummy is rejected as an overidentifying instrument (a test value of 5.65). The test statistic for the dummies nonwhite and smoke fall in the 10% critical region of the test (values greater than 2.71) but not in the 5% region. Correlations between these two variables and the regression error term are not immediate. For the union variable (where the case for rejection in light of the Sargan test is stronger) we may conceive of a positive correlation with transitory income: workers may enter unions if they believe these will promote their pay, and exit if they experience downfalls in their incomes (or alternatively if they lose their jobs).

The third line onwards in table 3 provide estimates of the intergenerational correlation for an increasingly larger set of instruments. In the third line for instance, the 0.73 estimate, we jointly test the validity of education and the south dummy (a Â^2 variate with two degrees of freedom), while in the next line the set of overidentifying instruments is extended to include the home ownership dummy. Of the estimates reported in the table, the only two speci…cations rejected at 5% are those that incorporate the union dummy (last two lines). Speci…cations that include the smoking dummy (without the union variable) are rejected at 10%, but not at 5%.

The range of estimates reported in the table (0.67 to 0.78) is substantially higher than those of Solon (1992). Solon's estimate of ¯ for family incomes using education as an instrument is equal to 0.53[12]. On the other hand,

[12] Also, Zimmerman (1992, table 10) using data from the American Nationonal Longi-

Behrman and Taubman (1990) whose study is also based on the PSID report estimates of ¯ range between 0.27 and 0.80. As mentioned earlier, we believe estimates of ¯ rise as children age. For instance, the …rst three estimates of table 3, taking the incomes of children …ve years earlier, are respectively 0.589 (0.146), 0.601(0.104) and 0.625(0.103). These are about 0.10 points less than the 1991 estimates.

C: MIMIC estimation results

In table 4 we estimate MIMIC speci…cations of intergenerational correlations in family incomes (panel A) and earnings of fathers and sons (panel B). The MIMIC framework may be used in the present context in order to shed light on three separate issues of concern: (i) the magnitude ¯ of the intergenerational correlation, (ii) the determinants of parental permanent income and (iii) the signal to total variance ratio. This latter quantity may be used directly to evaluate the bias of the ordinary least squares estimator of ¯. The Â^2 statistic (last column of the table) based on a comparison of the constrained model, i.e. the MIMIC estimate, with the unconstrained model, is used to test the null hypothesis that the MIMIC speci…cation is valid.

We estimate three distinct models for incomes and earnings. The …rst of these uses all instruments of table 3 with the exception of the union dummy (which was systematically rejected as an overidentifying instrument). The next set of estimates exclude the nonwhite and smoke dummies (whose related statistics fell in the 10% critical region of the Sargan test). Finally, we consider a parsimonious representation with only two instruments: the dummy for unskilled labour and the education of the family head.

Estimates of ¯ range between 0.71 and 0.77 for family incomes (panel A of table 4) and between 0.72 and 0.77 for earnings (panel B). We would have in fact expected that our earnings estimates of ¯ be smaller than those pertaining to family incomes, given the way our smaller fathers and sons sample was constructed. By looking at fathers' and sons' outcomes, we automatically discard children raised in female headed households, and children whose parents may have been unemployed, or who are themselves jobless, and accordingly do not report labour income. If persistence is more pronounced amongst disadvantaged groups than the middle classes[13], this would contribute to making family incomes-based estimates of ¯ higher than those

tudinal Survey presents instrumental variables estimates of ¯ for wages ranging between 0.38 and 0.71 when the Duncan index of socioeconomic status is used as an instrument.

[13]For instance, Zimmerman (1992) estimates transition matrices for the earnings of parents and children which are indicative of more persistence in the bottom and top

constructed from earnings data on fathers and sons. Our results however do not support this assumption.

Turning to the determinants of permanent income, we may note that the dummy variables unskill, ill, and nonwhite are all associated with negative e¤ects on permanent income. The coe¢cients on the dummies medins (medical insurance coverage) and house (home ownership) are positive. Estimates of these various magnitudes are statistically signi…cant for both the incomes and earnings results, and of the expected signs. However, we may note some qualitative di¤erences in the estimates depending on which of the two income concepts is used. In the case of earnings, the smoking dummy has a negative coe¢cient, but this estimate is not statistically di¤erent from zero (as indicated by the ± sign in the …rst line of the table). The south dummy has a negative, but statistically insigni…cant, coe¢cient in the top line of the income results, but smoking is estimated to have a positive e¤ect on permanent income. This positive coe¢cient on the smoking dummy in the income data comes to us as somewhat of a surprise, since tobacco consumption may be associated with an adverse e¤ect on health, and hence on the income generating capacity of individuals. It is possible however that this latter e¤ect only becomes apparent in the later stages of the life cycle of parents, but not so much when they are in their forties (as is the case in the present sample).

Turning to the returns to education, we have estimated a schooling coe¢cient in the order of 0.09 to 0.11 for family incomes, and in the order of 0.11 to 0.13 for earnings. These estimates are somewhat higher than the 6% to 8% consensual …ndings derived from standard speci…cations of earnings functions for the 1960s and early 1970s (see for instance the survey of Willis, 1986). There are however di¤erences in the sampling schemes, the de…nition of the dependent variables (the typical variable used in this literature is the hourly wage of the individual), the way schooling is measured, and more importantly, the choice of estimation method, which may all contribute to explain di¤erences between our …ndings and more specialized ones emanating from the returns to schooling literature.

The Â^2 speci…cation tests for the three alternative models of income and earnings continuities of table 4 do not reject the adequacy of the MIMIC framework. P-values for the income and earnings estimates (reported in the last column of the table) are respectively in the 0.13 to 0.69 and 0.67 to 0.82 range. In results not shown, the inclusion of the union dummy has however

quartile than for the middle income classes.

led to the rejection of the MIMIC specification on both the incomes and earnings data.

A final point we turn to in this results section is the estimation of the signal to total variance ratio. As discussed in our methods section, the parameter λ (equation 30) measures the bias of the ordinary least squares estimator. We have estimated λ to be in the order of 0.56 to 0.61 for total family incomes, and between 0.59 and 0.63 for earnings. Using an alternative covariance model, Zimmerman (1992) estimates λ to be 0.66 for earnings and 0.73 for hourly wages. Bowles (1972, table A1) reports estimates for λ ranging between 0.70 and 0.83 (for various income concepts).

Our figures imply that the bias of the ordinary least squares estimator is far from being negligeable. The OLS estimate corresponding to the MIMIC results of table 4 (the years 1967 and 1991) is equal to 0.433 for family incomes (table 2) and equal to 0.451 (with a standard error of 0.064) for the earnings of fathers and sons (results not shown). In order to obtain a corrected estimate of the intergenerational correlation we would have to multiply the OLS estimates by the inverse of the estimate of λ. On the basis of our MIMIC estimation results, we would therefore suggest to multiply $\hat{\rho}_{OLS}$ by a factor of 1.64 to 1.78 for family incomes, and a factor of 1.60 to 1.69 for earnings. In turn, these adjustments would imply slopes of 0.71 to 0.77 for family incomes, and 0.72 to 0.76 for earnings, which entail figures well within the range of our MIMIC estimations.

6 Conclusions

In this study we have proposed a framework for estimating the intergenerational correlation of incomes under the assumption that a model of the determinants of permanent income was known to the researcher, though this latter variable was unobserved. This has led us to propose two related estimators for the intergenerational correlation: a two-stage least squares procedure and a more efficient variant of the instrumental variables methodology–the MIMIC estimator. Following Mulligan (2000) and Han and Mulligan (1998), we have introduced uncertainty regarding the child's income in the family's resource allocation problem. The resulting Galtonian model of income transmission failed to satisfy the property of independence between regressor and disturbance at the level of the population. This has meant that, when moving from the population model to the measurement model, previously available probability limit formulae for OLS and related estimators were no longer valid. We have thus derived new expressions for the large sample biases of OLS and averaging methods. Likewise, we have proposed appropriate orthogonality requirements for the instrumental variables estimator to achieve consistency in the present form of the Galtonian model of income transmission.

We have examined continuities in the total family incomes of a US sample of parents and children extracted from the PSID. As a general rule, we have found that estimates of the intergenerational correlation rise as children age. For instance, an OLS regression of the child's 1987 income on that of her/his parents' 1967 resources estimates β at 0.30. Five years later, a regression of 1991 data on the same 1967 incomes of parents produces an estimate of 0.43 for the intergenerational correlation. Averaging the incomes of parents over time reduces the variance of the transitory component of income, and accordingly leads to an upward revision of the estimate of β. A regression of the child's 1987 income on a five-year average of parents' resources (1967 to 1971) estimates the intergenerational correlation at 0.38. However, looking at this same regression five years down the line for children, the estimate of β further increases to 0.50.

We have adopted various specifications for the model of the determinants of parental permanent income on the basis of which we have estimated the intergenerational correlation using two-stage least squares and MIMIC procedures. Our two-stage least squares estimates range between 0.67 and 0.78. These estimates are on the higher end of the spectrum of results emanating from the PSID and other US data sets. However, once again we have found that the same two-stage least squares procedures applied to various specifications with children's incomes observed in 1987 instead of 1991 reduced estimates of β by about 0.10, bringing them more in line with the findings from other studies which sampled earlier waves of the Panel.

We have also estimated MIMIC models for a subset of the specifications of the determinants of parental income retained in the two-stage least squares procedures. Our MIMIC estimates of the intergenerational correlation of family incomes are in the order of 0.71 to 0.77. These estimates of β are broadly similar to the corresponding two-stage least squares estimates. From the MIMIC results we have also constructed estimates of the signal to total variance ratio. Estimates of λ for family incomes range between 0.56 and 0.61, implying that the bias of the ordinary least squares estimator may be in the order of 40%. We have also repeated the same MIMIC estimations on a smaller sample of fathers and sons who report non-zero annual earnings.

MIMIC estimates of the intergenerational correlation ranged between 0.72 and 0.77, and λ was estimated to lie in the range of 0.59 to 0.63.

We may note that the MIMIC framework we have proposed provides a more general representation of the relation between a child's income and her parents' characteristics. For instance, the results of table 4 show that the child's income is increasing in the education of the parent family head, and that growing up in a family with access to medical insurance also has a positive effect on adult income. The magnitude of the marginal effect of each of these variables may simply be obtained upon multiplying the estimate of β by the estimate of the respective coefficient on the family background variable. A representation of the relation between a child's income and a set of socio-economic characteristics of his parents may provide a useful framework for thinking about income maintenance policies in a long run perspective, alongside the estimates derived from the usual Galtonian regressions.

The linear representation of the association between the incomes of parents and children, the Galtonian framework, can only be traced back to a family resource allocation problem to the extent that resource constraints are inexistent, or alternatively that credit markets for human capital loans are perfect. Outside of this, the slope estimate of the Galtonian regression is no longer a consistent estimator of the correlation between the abilities of parents and children. The simulations of Han and Mulligan (1998) do not treat the problem of measurement error in permanent incomes, so that it is still not known how errors-in-variables biases interact with specification biases (arising from the use of linear functional forms in presence of binding wealth constraints) to arrive at an estimate of β.

Relaxing the linearity assumption in empirical work is certainly a task worth undertaking. One potential research direction may consist in estimating a model of switching regressions. We assume the existence of separate linear models of income transmission for wealth constrained and unconstrained families. We do not know which families are wealth constrained, however we specify a model for the probability of being in either of the two regimes. Because of the usual errors in the measurement of permanent income, the model must most probably be embedded within a simultaneous equations framework (Maddala, 1983; chs. 7 and 8). Such an approach would still come short of providing direct estimates of the structural parameters of a family utility model with wealth constraints. Nonetheless, it would begin to address empirically the question as to how wealth constraints affect the process of income inheritance when it is not known which families face resource constraints and permanent income is also subject to errors of measurement.

7 References

Abul Naga R. and J. Krishnakumar (1999): "Panel Data Estimation of the Intergenerational Correlation of Incomes", discussion paper no. 9910, DEEP, Université de Lausanne.

Altonji J. and T. Dunn (1991): "Relationships Among the Family Incomes and Labor Market Outcomes of Relatives", Research in Labor Economics 12, 263-310.

Becker G. and N. Tomes (1986): "Human Capital and the Rise and Fall of Families", Journal of Labor Economics 4, S1-S39.

Behrman J. (1997): "Intrahousehold Distribution and the Family",in Rosenzweig M. and O. Stark eds: Handbook of Population and Family Economics, Amsterdam Elsevier.

Behrman J. , R. Pollak and P. Taubman (1982): "Parental Preferences and Provision for Progeny", Journal of Political Economy 90, 52-73.

Behrman J. , R. Pollak and P. Taubman (1995): "The Wealth Model: E¢ciency in Education and Equity in the Family", in Behrman J. , R. Pollak and P. Taubman eds: From Parent to Child: Intra-Household Allocations and Intergenerational Relations in the United States, Chicago, University of Chicago Press.

Behrman J. and P. Taubman (1990):"The Intergenerational Correlation Between Children's Adult Earnings and Their Parents' Incomes: Results from the Michigan Panel Study of Income Dynamics", Review of Income and Wealth, 36, 115-127.

Björklund A. and M. Jantti (1997): "Intergenerational Income 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.

Godfrey L. (1988): Misspeci…cation Tests in Econometrics, Cambridge, Cambridge University Press.

Griliches Z. (1986): "Economic Data Issues", in Griliches Z. and M. Intriligator eds: Handbook of Econometrics volume 3, Amsterdam, North Holland.

Han S. and C. Mulligan (1998): "Human Capital, Heterogeneity, and the Estimation of Degrees of Intergenerational Mobility", mimeo, University of Chicago.

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

Jöreskog K. and A. Goldberger (1975): "Estimation of a Model with Multiple Indicators and Multiple Causes on a Single Latent Variable", Journal of the American Statistical Association 70, 631-639.

Klepper S. and E. Leamer (1984): "Consistent Sets of Estimates for Regressions with Errors in All Variables", Econometrica 52, 163-183.

Maddala G. (1983): Limited-Dependent and Qualitative Variables in Econometrics, Cambridge, Cambridge University Press.

Mulligan C. (2000): "Galton versus Human Capital Approach to Inheritance", Journal of Political Economy, forthcoming.

Reville R. (1995): "Intertemporal and Life Cycle Variation in Measured Intergenerational Earnings Mobility", mimeo.

Solon G. (1989): "Biases in the Estimation Intergenerational Earnings Correlations", 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.

Willis R. (1986): "Wage Determinants: A Survey and Reinterpretation of Human Capital Earnings Functions", in Ashenfelter O. and R. Layard eds: Handbook of Labor Economics, volume1, Amsterdam, Elsevier.

Zimmerman D. (1992): "Regression Toward Mediocrity in Economic Stature", American Economic Review 82, 409-429.

Table 1 : Descriptive statistics

variable	mean	coeff. variation
Age of parent head in 1968	43.68	0.20
Age of child head in 1992	39.73	0.21
Parents' 1967 family income	10,767	0.59
Child's 1991 family income	12,778	1.17
Non-whites (%)	10.98	--
Schooling interval of parent head	12 years	--
Unskilled parent head (%)	8.95	--

Note: The income concept is total family income measured in 1967 dollars.

Table 2 : Estimation by OLS and the method of averaging

Variables	1967	1967-68	1967-69	1967-70	1967-71	1971
1987	0.297 (0.0053)	0.342 (0.054)	0.360 (0.055)	0.375 (0.055)	0.379 (0.056)	0.295 (0.049)
1991	0.433 (0.057)	0.467 (0.058)	0.478 (0.058)	0.502 (0.059)	0.501 (0.059)	0.376 (0.053)
1987-91	0.365 (0.046)	0.403 (0.047)	0.416 (0.047)	0.432 (0.048)	0.433 (0.048)	0.333 (0.047)

Notes

1 The income concept is total family income measured in 1967 dollars.

2 1987-91 signifies that the child's income is averaged over the corresponding five-year period; 1967-68 is a two-year average of parents' resources, etc.

3 Standard errors are reported inside parentheses.

Table 3 : Instrumental variables estimation results

b	unskill	edu	south	house	medins	ill	nonwhite	smoke	union	Sargan test
0.667 (0.154)	x									--
0.710 (0.110)	x	x								0.165
0.733 (0.109)	x	x	x							1.300
0.770 (0.106)	x	x	x	x						3.152
0.780 (0.105)	x	x	x	x	x					3.702
0.751 (0.103)	x	x	x	x	x	x				7.588
0.775 (0.102)	x	x	x	x	x	x	x			9.186
0.728 (0.728)	x	x	x	x	x	x		x		10.717
0.752 (0.100)	x	x	x	x	x	x	x	x		12.342
0.690 (0.101)	x	x	x	x	x	x			x	17.732*
0.694 (0.098)	x	x	x	x	x	x	x	x	x	23.174*

Notes

1 The income concept is total family income measured in 1967 dollars.
2 The parents' income pertains to 1967 and the child's income to 1991.
3 An **x** mark indicates that the corresponding variable is included in the set of instruments.
4 A * indicates that the Sargan test of over-identifying restrictions rejects the corresponding specification at a 5% probability of type-I error.

Table 4 : MIMIC estimation results

A : Family incomes of parents (1967) and children (1991)

b	**unskill**	**edu**	**south**	**house**	**medins**	**ill**	**nonwhite**	**smoke**	**l**	**Khi sq. test** [P-value]
0.770 (0.100)	-0.387 (0.071)	0.088 (0.010)	-0.076° (0.044)	0.190 (0.046)	0.104 (0.048)	-0.123 (0.055)	-0.274 (0.066)	0.079 (0.040)	0.563	11.32 [0.125]
0.763 (0.103)	-0.446 (0.071)	0.089 (0.010)	-0.128 (0.043)	0.206 (0.047)	0.114 (0.049)	-0.117 (0.056)			0.567	7.10 [0.214]
0.711 (0.110)	-0.539 (0.072)	0.111 (0.010)							0.609	0.16 [0.686]

B : Earnings of fathers (1967) and sons (1991)

b	**unskill**	**edu**	**south**	**house**	**medins**	**ill**	**nonwhite**	**smoke**	**l**	**Khi sq. test** [P-value]
0.721 (0.106)	-0.189 (0.095)	0.114 (0.012)	-0.117 (0.053)	0.124 (0.055)	0.130 (0.062)	-0.185 (0.069)	-0.306 (0.092)	-0.041° (0.044)	0.626	3.69 [0.815]
0.730 (0.109)	-0.255 (0.095)	0.118 (0.012)	-0.169 (0.051)	0.140 (0.055)	0.129 (0.063)	-0.181 (0.070)			0.619	2.72 [0.743]
0.766 (0.119)	-0.314 (0.096)	0.136 (0.012)							0.590	0.18 [0.669]

Notes

1 n = 592 for the income data; n = 369 for the earnings data.

2 A ° mark indicates that the corresponding coefficient is not statistically different from zero at the 5% significance level.