

Biases of the ordinary least squares and instrumental variables estimators of the intergenerational earnings elasticity: Revisited in the light of panel data

Ramses H. Abul Naga

Received: 1 February 2007 / Accepted: 7 August 2007 / Published online: 19 October 2007
© Springer Science + Business Media B.V. 2007

Abstract The OLS estimator of the intergenerational earnings correlation is biased towards zero, while the instrumental variables estimator is biased upwards. The first of these results arises because of measurement error, while the latter rests on the presumption that the education of the parent family is an invalid instrument. We propose a panel data framework for quantifying the asymptotic biases of these estimators, as well as a mis-specification test for the IV estimator.

Keywords Intergenerational mobility · Measurement error · Panel data · Biased estimation

JEL Classification I3 · J62

1 Introduction

Researchers often agree that the level of the intergenerational earnings elasticity is a useful benchmark for measuring the degree of attainment in an equal opportunities social justice objective. There is more debate however and less convergence, regarding the opinions held on the process through which income status is transmitted from parent to child. According to [8] for instance, parents determine the level of human capital they wish to invest in children, depending on their offsprings' levels of cultural and learning attributes, known as endowments. If credit markets are perfect, the intergenerational elasticity of earnings reflects the inheritance of endowments between parents and children. There also exists a literature where family optimizing behaviour does not feature prominently, yet also yielding important predictions about the mechanisms of income transmission [12, 17].

R. H. Abul Naga (✉)
IEMS, Université de Lausanne, CH1015, Lausanne, Switzerland
e-mail: rabulnag@unil.ch

The last decade has certainly witnessed a renewed interest in questions pertaining to the distribution of income at one point in time, and also in the extent to which children inherit the economic status of their parents. In the empirical intergenerational mobility literature, researchers have had to deal with the challenging question of estimating the elasticity between life-time incomes of parents and children using short time series, and at times a unique observation, on family resources. It has been immediately recognized in early contributions to the literature [7, 11] that the use of annual income measures to proxy permanent incomes produced an errors in variables problem, and as a result, the ordinary least squares estimator of the intergenerational earnings elasticity was biased towards zero.

As a response to this problem, researchers working with panel data have averaged the resources of the parent family (i.e. the explanatory variable) over typically three to five years in order to reduce the bias resulting from the estimation of the standard relation linking the incomes of parents and children, known as the Galtonian regression model. Illustrations of averaging methods can be found in Behrman and Taubman [9], Solon [28], Zimmerman [30], Björklund and Jantti [10], Mulligan [26] and others. A problem remaining though, is that not all data sets provide repeated observations on the parent family's resources. A typical example in this context is the UK's National Child Development Study (see for instance [13] for a discussion). When no time series variation is available on the parent family's resources, it is typical then to instrument the unique measurement on this variable using the education of the family head. A standard argument formalized by Solon [28] is that when the parent family head's education features as an explanatory variable in a model of the determinants of the child's income, the instrumental variables estimator of the intergenerational elasticity is biased away from zero.

Thus it has become a known result that the ordinary least squares and instrumental variables estimators bracket the intergenerational elasticity.¹ There have been several efforts in quantifying the bias of the ordinary least squares estimator since the work of Bowles [11], and estimates provided by Zimmerman [30], Abul Naga [1] and Mazumder [24] suggest that the extent of the resulting bias is perhaps in the order of 30% or more. Little is known however about the degree to which the instrumental variables methodology over-estimates the intergenerational elasticity. A key quantity in evaluating the bias of the ordinary least squares estimator is the variance ratio of permanent to total income. A remaining question then is what comparable parameters (for which, prior knowledge of their magnitudes is required) play a similar role in the correction of the bias of the instrumental variables estimator.

There are two reasons why we feel that this exercise may be of direct relevance to researchers and policy makers. Firstly, having an order of magnitude about the bias of the instrumental variables estimator provides another route to refining our knowledge on the extent of income continuities across generations. More importantly though, if it is found that the bias of the instrumental variables estimator is small, or negligible,² earlier results which may have been read to exaggerate the underlying intensity of income inheritance in the population, may be re-appraised in a different light.

¹This result is often emphasized in the literature. See for instance Björklund and Jantti [10] and Dearden et al. [13].

²Solon [28], referring to related empirical evidence from the United States, entertains this hypothesis.

Our starting point in this paper is a structural model of income transmission where the child's permanent income is explained by the parent family's permanent income and education. We shall then be testing the assumption that the structural model is under-identified in the context of cross-section data where a unique noisy measurement is available on the parent family's permanent income. Essentially, this approach will amount to testing the *null hypothesis* that the parent family's education is a valid instrumental variable. The difficulty with such a task is that under the *alternative hypothesis* (i.e. when education is not a valid instrument) the structural model is no longer identified in the cross-section environment. Thus one is inevitably led to formulate a more general model which remains identified *under both the null and alternative hypotheses*.

Aigner et al. [3] point to three major directions in trying to overcome under-identification. Firstly, one may specify an additional structural equation, from which one or several over-identifying restrictions may be sought. This approach is systematically developed in the context of the simultaneous equations literature. A second approach is to assume the existence of multiple equations yielding alternative indicators on an error-ridden variable. Clearly, panel data provide a specific context of this second approach. Thirdly, the specification of a dynamic structure for the variable measured subject to noise may increase the scope for identification.

Here we shall pursue successively the second and third approaches to resolving the under-identification problem. Firstly, in order to evaluate the biases of the ordinary least squares and instrumental variables estimators we use panel data to derive consistent estimators for economic relations observed subject to measurement error.³ As a by-product of our discussion therefore, we also propose a consistent estimator of the intergenerational elasticity when repeated measurements are available on the parent family's income. Likewise, a test for evaluating the claim that the instrumental variables estimator of the intergenerational elasticity is upwardly inconsistent becomes feasible in the panel data context where the model remains generally identified under both the null and alternative assumptions of interest. A test of this hypothesis, which exploits the information on the time series variation in earnings, has not been proposed to date in the literature.

Secondly, we show that when we part with the assumption that the permanent income of the parent family is a time-invariant quantity, the specification of a dynamic structure for this variable allows us to identify a more general class of income transmission models. Such models allow for correlated heterogeneity and the presence of child-specific permanent income trends. These developments form the basis of a later section of the paper devoted to the elaboration of various research directions.

In order to analyze the large sample biases of the ordinary least squares [OLS] and instrumental variables [IV] estimator in a joint framework, we express these as functions of the structural parameters of a model of income transmission. We then inquire about the nature of the data required in order to recover these unknown structural parameters from the biases of the two estimators. It turns out that the two fundamental quantities required in order to perform this exercise are (1) the correlation between education and permanent income and (2) the variance ratio of

³As shown by Griliches and Hausman [19], Hsiao and Taylor [23] and others, a wide range of errors in variables models are identifiable in the panel data context.

permanent to total income. In turn, we propose estimators for these quantities when panel data are available to the researcher. The proposed consistent estimator of the intergenerational elasticity and test for the instrumental variables estimator also arise from our study of the identification of the structural model of income transmission.

We then use the panel study of income dynamics, a longitudinal survey from the United States, in order to evaluate numerically the biases of the OLS and IV estimators. Our results suggest that the first of these under-estimates the intergenerational elasticity by about 20 to 40%, while the latter over-states it by 1 to 32%.

In the next section we examine the identification of a structural model of income transmission initially suggested by Solon [28], where the OLS estimator is biased towards zero, and IV is biased away from zero. In Section 3 we discuss in what ways panel data may allow the researcher to identify the structural parameters of a model of income inheritance and to test for mis-specification of the instrumental variables estimator. In Section 4 we provide a brief description of our data. In Section 5 we present estimates of the biases resulting from these two methods using our US data.

Section 6 is devoted to the elaboration of various research directions. In particular, we consider more general specifications for the income transmission model, inspired from the micro-econometrics panel data literature (e.g. [5]). There, we show that in presence of correlated heterogeneity⁴ the rule stating that the OLS and IV estimators bracket the intergenerational elasticity need no longer hold. We also discuss what type of processes are required to govern the time-series evolution of income in order to identify the structural model under these alternative assumptions. Section 7 concludes the paper. Three technical appendices to the paper are available from the author upon request.

2 Biases of OLS and IV

In this section we study the biases of the ordinary least squares and instrumental variables estimator of the intergenerational elasticity. We show that without repeated measurements on parental income, these biases cannot generally be quantified.

The framework we consider is a Galtonian regression of the child's long-run income, $\log(I_c)$ on that of her/his parents', $\log(I_p)$, where $\log(\cdot)$ is the logarithmic function. Let $\eta_c = \log(I_c) - E(\log(I_c))$ and $\eta_p = \log(I_p) - E(\log(I_p))$. η_c and η_p are the logarithms of the long-run incomes of children and parents, measured in deviation from their respective means, which we shall denote below in more simple terms as permanent incomes. The Galtonian regression model may be written as:

$$\eta_{ci} = \beta\eta_{pi} + \zeta_i \quad (1)$$

where $i = 1, \dots, n$ indexes data on family i , and it is in general assumed that $E(\zeta_i|\eta_{pi}) = 0$ for all i . The interpretation of long-run income as a time-invariant quantity follows a long tradition in the estimation of the Galtonian model. In Section 6 however we shall return to this assumption and to the assumption that $E(\zeta_i|\eta_{pi}) = 0$.

⁴By correlated heterogeneity it is meant that in the structural model of income inheritance the disturbance term is correlated with the explanatory variables.

In contrast with the theoretical model 1, it is assumed that the researcher observes annual data y_{it} and x_{is} which are taken as noisy measurements on the permanent incomes η_{ci} and η_{pi} :

$$y_{it} = \eta_{ci} + v_{it} \quad (2)$$

$$x_{is} = \eta_{pi} + u_{is} \quad (3)$$

It is to be noted that most often in practice $t \neq s$. In the context of the Panel Study of Income Dynamics for instance it has become common to regress the child's 1990s income on that of the parent family in the late 1960s. This practice is motivated by an attempt to measure earnings of parents and children at a similar stage of the life-cycle (around the age of 40). There it is believed that the magnitude of the earnings transitory variance component is considerably reduced in comparison to its level in the early years of the life-cycle.

We are working in the context of a micro panel where n is large relative to the time dimension. All expectations below therefore pertain to the cross-section. Define e^* as the education of the parent household head and $e = e^* - E(e^*)$, as education measured in deviation from its mean. We shall make the following assumptions regarding errors of measurement, for all individuals i , and for all time periods s and t :

$$E(v_{it}\eta_{ci}) = E(u_{is}\eta_{pi}) = 0 \quad (4)$$

$$E(v_{it}e_i) = E(u_{is}e_i) = 0 \quad (5)$$

$$E(v_{it}\eta_{pi}) = E(u_{is}\eta_{ci}) = 0 \quad (6)$$

$$E(v_{it}u_{is}) = 0 \quad (7)$$

That is, measurement errors v and u are mutually uncorrelated, and are also uncorrelated with the permanent incomes of children and parents as well as the education of the parent family.

Below we shall assume that all data are measured in deviation from their cross-sectional means. The researcher regresses y_{it} on x_{is} by ordinary least squares to obtain an estimator $\hat{\beta}_{OLS}$:

$$\hat{\beta}_{OLS} = \frac{\sum_i y_{it}x_{is}}{\sum_i x_{is}^2} \quad (8)$$

Because x_{is} is a noisy measurement on η_{pi} , the above estimator is biased towards zero. An alternative estimation strategy for the Galtonian model 1 consists in instrumenting x_{is} using e_i . This produces the following instrumental variables estimator:

$$\hat{\beta}_{IV} = \frac{\sum_i e_i y_{it}}{\sum_i e_i x_{is}} \quad (9)$$

In order to formalize the argument that the instrumental variables estimator may be upwardly biased, we shall assume, following Solon [28], that parental education features as an explanatory variable in a structural model of the determinants of the child's income:

$$\eta_{ci} = \delta_0 \eta_{pi} + \delta_1 e_i + v_i : \quad \delta_1 \geq 0 \quad (10)$$

where v_i is taken to be uncorrelated with e_i and η_{pi} :⁵

$$E(v_i e_i) = 0 \quad \text{and} \quad E(v_i \eta_{pi}) = 0 \tag{11}$$

It is to be noted that education is not the only instrument that has been used in practice. For instance, in Zimmerman [30], Fortin and Lefebvre [15] and Oreopolous [27] variables such as socioeconomic status, occupational status or neighborhood status perform the role of identifying instruments. Also, one important issue pointed out by Ermisch and Francesconi [14] is that the proposed instrument itself may be subject to measurement error. Furthermore, in the structural model 10–11 there may be several, not just one variable alongside η_p featuring as determinants of η_c . These two extensions of the above structural model are discussed in turn in Appendix 2 of the paper.

Returning to the model 10–11, from Eq. 1 we have that $\beta = E(\eta_p \eta_c) / E(\eta_p^2)$. Hence, it also follows from Eq. 10 that $E(\eta_p \eta_c) = \delta_0 E(\eta_p^2) + \delta_1 E(\eta_p e)$, i.e. that

$$\beta = \delta_0 + \delta_1 \frac{E(\eta_p e)}{E(\eta_p^2)} \tag{12}$$

In what follows, we shall refer to β as the *intergenerational elasticity*, and δ_0 as the *intergenerational slope* coefficient. Also, while we refer to Eq. 1 as the Galtonian model, we shall call the relation 10 the *structural model of income transmission*. Clearly, these definitions coincide when $\delta_1 = 0$.

Returning to Eq. 12, Solon shows that if parental education is a determinant of the child’s income, $\hat{\beta}_{IV}$ and $\hat{\beta}_{OLS}$ respectively provide upper and lower bound estimates of the intergenerational elasticity β , in the sense that

$$p \lim(\hat{\beta}_{OLS}) < \beta < p \lim(\hat{\beta}_{IV}) \tag{13}$$

where $p \lim(\cdot)$ is the probability limit operator. Probability limit formulas for these two estimators are given in Solon [28]. Here we shall write them in slightly different forms, as functions of the parameters of the structural model 10:

$$p \lim(\hat{\beta}_{OLS}) = \frac{\delta_0 E(\eta_p^2) + \delta_1 E(\eta_p e)}{E(\eta_p^2) + E(u^2)} \tag{14}$$

$$p \lim(\hat{\beta}_{IV}) = \delta_0 + \delta_1 \frac{E(e^2)}{E(\eta_p e)} \tag{15}$$

By defining $\lambda = E(\eta_p^2) / [E(\eta_p^2) + E(u^2)]$ as the variance ratio of permanent to total income, also known as the *signal to total variance ratio*, we may use Eq. 12 in order to obtain a more familiar expression for $p \lim(\hat{\beta}_{OLS})$:

$$p \lim(\hat{\beta}_{OLS}) = \beta \lambda \tag{16}$$

from which it comes out more clearly that $\hat{\beta}_{OLS}$ is asymptotically biased downwards, as the variance ratio λ is smaller than one. The instrumental variables estimator on the other hand is biased away from zero provided $\delta_1 > 0$ and education is positively

⁵We shall relax this assumption in Section 6.

correlated with permanent income. Define $\rho = E(\eta_p e) / [E(\eta_p^2) E(e^2)]^{1/2}$ as the correlation coefficient between permanent income and education. The following result is a direct consequence of Eq. 15:

Proposition 1 *Let the structural model mapping the permanent incomes of parents and children be given by Eq. 10–11 where $\delta_1 > 0$. Then the large sample bias of $\hat{\beta}_{IV}$ is decreasing in the correlation ρ between education and permanent income; the bias vanishes as $\rho \rightarrow 1$.*

Solving the above system 14 and 15, it may be verified that δ_0 and δ_1 , the parameters of Eq. 10, relate to the population second moments in the following manner

$$\delta_0 = \frac{p \lim(\hat{\beta}_{OLS})/\lambda - \rho^2 p \lim(\hat{\beta}_{IV})}{1 - \rho^2} \quad (17)$$

$$\delta_1 = \frac{[p \lim(\hat{\beta}_{IV}) - p \lim(\hat{\beta}_{OLS})/\lambda] E(\eta_p e) / E(e^2)}{1 - \rho^2} \quad (18)$$

That is, because a single measurement x_{is} will not identify the variance of permanent income, it is not possible to evaluate λ or ρ . From this observation it also follows that under such circumstances it is not possible either to identify the structural model 10, nor is it possible to evaluate the large sample biases of the OLS and IV estimators.

3 Panel data and measurement error

With panel data however, we show that the process of income transmission is identified; hence we may readily derive formulas to evaluate numerically the biases of these estimators. Let x_{is} and $x_{i\tau}$ be two measurements on family income for periods s and τ . Zimmerman [30] assumes that the measurement error exhibits an autoregressive structure of order one, so that the correlation between u_{is} and $u_{i\tau}$ in Eq. 3 diminishes as the time distance between periods s and τ increases. On the other hand, Altonji and Dunn [4] and Abul Naga and Krishnakumar [2] suggest that errors of measurement exhibit a moving average structure. That is, passed a certain time period – most probably two years according to their respective findings – changes in annual earnings are serially uncorrelated (see Table 1). The study of Mazumder [24] departs from the three studies discussed above in several respects, two of which are highlighted in the Table. Firstly, parental permanent income rather than being treated as a time-invariant fixed effect, is allowed to vary over time and is modelled as a random walk. Secondly, a distinction is drawn between measurement error and transitory earnings components, and the variance of the latter is allowed to vary over time. Finally, Haider and Solon [20] while treating permanent income as a fixed effect, model measurement error as being mean-reverting, a process considerably different from classical errors in variables.

We shall return to the Mazumder and Haider and Solon contributions in Section 6. For now, we note that it is possible to generalize the approaches of the first three

Table 1 Various models of the permanent and transitory/measurement error earnings components used in the intergenerational mobility literature

| Study | Permanent Component | Transitory / measurement error component |
|--------------------------------|-----------------------------|---|
| Altonji and Dunn [4] | Time invariant fixed effect | $MA(2)$ |
| Zimmerman [30] | Time invariant fixed effect | $AR(1)$ |
| Abul Naga and Krishnakumar [2] | Time invariant fixed effect | $MA(2)$ |
| Mazumder [24] | Random walk | Transitory component separate from measurement error Transitory variance component changes over the life cycle |
| Haider and Solon [20] | Time invariant fixed effect | Mean reverting |

studies surveyed in Table 1 by considering an *ARMA* specification for the error process. Consider an *ARMA*(1, q) form:

$$u_{is} = \phi u_{is-1} + \varepsilon_{is} + \theta_1 \varepsilon_{is-1} + \dots + \theta_q \varepsilon_{is-q} \tag{19}$$

The case considered by Zimmerman [30] sets $\theta_\tau = 0$ for all τ , while Abul Naga and Krishnakumar [2] set $\phi = 0$, and $\theta_\tau = 0$ for all $\tau > 2$.

Now consider the identification of $E(\eta_p^2)$ in the light of Eq. 19. For this purpose, it is useful to consider moments of the type $E[(x_{it} - \phi x_{it-1})x_{is}]$. Noting that $\phi x_{it-1} = \phi \eta_{pi} + \phi u_{it-1}$, we have that

$$x_{it} - \phi x_{it-1} = (1 - \phi)\eta_{pi} + \varepsilon_{it} + \theta_1 \varepsilon_{it-1} + \dots + \theta_q \varepsilon_{it-q} \tag{20}$$

That is, the above quasi-difference is an *MA*(q) process. Furthermore,

$$E[(x_{it} - \phi x_{it-1})x_{is}] = E\left[\left((1 - \phi)\eta_{pi} + \varepsilon_{it} + \theta_1 \varepsilon_{it-1} + \dots + \theta_q \varepsilon_{it-q}\right) (\eta_{pi} + u_{is})\right] \tag{21}$$

Provided therefore that $\tau - q > s$, i.e. $\tau > s + q$, the above ε terms are uncorrelated with u_{is} and Eq. 21 simplifies to $(1 - \phi)E(\eta_p^2)$. Accordingly, with an *ARMA*(1, q) specification, and provided $\phi \neq 1$,⁶ we can identify $E(\eta_p^2)$ using the moment condition

$$E(\eta_p^2) = \frac{E[(x_{it} - \phi x_{it-1})x_{is}]}{1 - \phi} \quad \tau > s + q \tag{22}$$

It is instructive to consider how the above moment condition simplifies in presence of the *AR*(1) and *MA*(2) specifications discussed above. For the auto-regressive specification the terms ε_{it-1} to ε_{it-q} vanish from Eq. 21, so that Eq. 22 becomes valid subject to the weaker requirement $\tau > s$. Under the moving average specification $\phi = 0$, so that Eq. 22 reduces to $E(\eta_p^2) = E(x_{it}x_{is})$, with $|\tau - s| > q$ for a general *MA*(q) process.⁷

⁶The case $\phi = 1$ (known in the literature as the *unit-root model*) is taken up in Section 6.

⁷That is, in this case both lags and leads of x_{is} may be used to identify $E(\eta_p^2)$.

With this extra condition we may note that it is now possible to identify the structural model Eq. 10 for the determinants of the child's income. Write $m(x, y)$ as the sample covariance between variables x and y , and let $m[(x_{it} - \phi x_{it-1})x_{is}]$ denote the ratio of Eq. 22. Let

$$\hat{\lambda} = \frac{m[(x_{it} - \hat{\phi}x_{it-1})x_{is}]}{m(x_{is}^2)} \quad (23)$$

$$\hat{\rho} = \frac{m(x_{is}e_i)}{[m[(x_{it} - \hat{\phi}x_{it-1})x_{is}]m(e_i^2)]^{1/2}} \quad (24)$$

denote estimators for λ and ρ , where $\hat{\phi}$ is any consistent estimator of ϕ .⁸ $\hat{\lambda}$ is consistent for λ under condition Eq. 22,⁹ while $\hat{\rho}$ is consistent under the additional assumption that transitory income is uncorrelated with educational attainment.¹⁰ Note finally that any estimator $\hat{\alpha}$ consistently estimates $p \lim(\hat{\alpha})$ so that we may consistently estimate δ_0 and δ_1 using the following expressions:

$$\hat{\delta}_0 = \frac{\hat{\beta}_{OLS}/\hat{\lambda} - \hat{\rho}^2\hat{\beta}_{IV}}{1 - \hat{\rho}^2} \quad (25)$$

$$\hat{\delta}_1 = \frac{[\hat{\beta}_{IV} - \hat{\beta}_{OLS}/\hat{\lambda}]m(x_{is}e_i)/m(e_i^2)}{1 - \hat{\rho}^2} \quad (26)$$

We may note that Eq. 26 provides a direct test for the null hypothesis that $\hat{\beta}_{IV}$ is consistent, against the alternative that it is upwardly inconsistent. If δ_1 is estimated to be statistically *not* different from zero, it may be concluded that $p \lim(\hat{\beta}_{IV}) = p \lim(\hat{\beta}_{OLS}/\hat{\lambda}) = \beta$, i.e. that the instrumental variables estimator is consistent. Under the alternative, taking probability limits in Eq. 26,¹¹ we obtain the following result:

Proposition 2 *Let the structural model mapping the permanent incomes of parents and children be given by Eq. 10–11 and assume $\rho < 1$. Then $\delta_1 > 0$ implies that $p \lim(\hat{\beta}_{IV}) > \beta$.*

By going back to the structural model 10 we obtain an alternative intuition for this test. The argument is as follows: if $\delta_1 = 0$, this implies that parental education is not a determinant of the child's income, and hence the instrumental variables estimator achieves consistency. An evaluation of the hypothesis that $\hat{\beta}_{IV}$ is consistent is therefore undertaken here by treating Eq. 10 as an extended regression of Eq. 1, where it is maintained, under the null hypothesis, that $\delta_1 = 0$.

⁸An Anderson–Hsiao procedure is used here to estimate ϕ . See Appendix 1 for further detail.

⁹When several moment conditions of the type Eq. 22 are available, say m_1 and m_2 , $\hat{\lambda}$ is constructed as an average of these, say $(m_1 + m_2)/2m(x_{is}^2)$; more on this in Section 5 below.

¹⁰This statement follows from the fact that under the assumption $E(u_{is}e_i) = 0$, we have that $E(x_{is}e_i) = E[(\eta_{pi} + u_{is})e_i] = E(\eta_{pe})$.

¹¹Alternatively, returning to Eq. 18.

Substituting Eqs. 25 and 26 into Eqs. 12 – alternatively, using Eq. 16 – entails the following estimator for β :

$$b = \hat{\beta}_{OLS} / \hat{\lambda} \tag{27}$$

which is a feasible form for the *consistent adjusted least squares* estimator [25].

In our empirical applications we shall evaluate the large sample biases of the OLS and IV estimators in percentage terms, i.e. in the form $\lambda = p \lim(\hat{\beta}_{OLS})/\beta$ and $\gamma = p \lim(\hat{\beta}_{IV})/\beta$. In the sample these may be evaluated by the quantities Eq. 23 and

$$\hat{\gamma} = \hat{\lambda} \frac{\hat{\beta}_{IV}}{\hat{\beta}_{OLS}} \tag{28}$$

The above equation highlights one important relation between the calculated biases of the two estimators: the higher λ is (viz. the smaller the bias of OLS), the higher the bias γ imputed to $\hat{\beta}_{IV}$. One important question therefore for the data analyst is whether the choice of an error process may affect the biases of the two estimators in any systematic direction. Let $\lambda(\phi, \theta_1, \theta_2)$ denote the signal to total variance ratio underlying an *ARMA*(1, 2) error process in relation to Eq. 19. Using this notation, define $\lambda(\phi, \theta_1, 0)$, $\lambda(\phi, 0, 0)$ and $\lambda(0, \theta_1, \theta_2)$ respectively as the signal to total variance ratio underlying an *ARMA*(1, 1), an *AR*(1) and an *MA*(2) error process. The following proposition states sufficient conditions on the structural parameters ϕ, θ_1 and θ_2 for the signal to total variance ratio to increase in moving from an *ARMA*(1, 2) error process to each of the three nested error processes stated above:¹²

Proposition 3 *Let the relation between observed income and permanent income be given by Eqs. 3 and 19 where $E(\eta_{pi}u_{is}) = 0$, ε_{is} are iid disturbances and $q = 2$. Then,*

- (1) $\phi(\phi + \theta_1)\theta_2 \geq 0$ entails $\lambda(\phi, \theta_1, \theta_2) \leq \lambda(\phi, \theta_1, 0)$
- (2) $\phi[(\phi + \theta_1)\theta_2 + \theta_1] \geq 0$ entails $\lambda(\phi, \theta_1, \theta_2) \leq \lambda(0, \theta_1, \theta_2)$
- (3) $\phi[(\phi + \theta_1)\theta_2 + \theta_1] \geq 0$ also entails $\lambda(\phi, \theta_1, \theta_2) \leq \lambda(\phi, 0, 0)$

The proposition states that for nested specifications of the error process, conditions may be met so that the signal to total variance ratio is no higher in the context of the more general specification.¹³ On the other hand, such comparisons cannot be made in the context of non-nested specifications such as an *AR*(1) versus an *MA*(2) error process. There, the researcher cannot anticipate which of the two specifications will entail a higher λ .

In Section 5 we use data from the panel study of income dynamics in order to estimate the biases of these two estimators of the intergenerational elasticity parameter β in relation to various error processes nested within the *ARMA* family.

¹²See Appendix 1 for a proof of the proposition. The results are limited here within the confines of an *ARMA*(1, 2) error process as we shall not exceed this level of generality in our applications of Section 5. Comparisons however may also generally be made between nested specifications within the context of an *ARMA*(1, q) error process.

¹³Note that these conditions are not necessary: the bias of OLS may fall in moving from an *ARMA*(1, 2) to an *MA*(2) specification, yet the conditions (2) of the proposition may not be met. Note however that often in practice $\theta_1 \leq 0$, $\theta_2 \geq 0$ and $|\theta_1| < |\phi|$, so that these conditions are quite likely to be satisfied in relation to income processes.

4 Data

In order to quantify the biases of the ordinary least squares and instrumental variables estimators of the intergenerational elasticity, we look at earnings continuities in a US sample of parents and children. Our data are extracted from the SRC file, the random sample, of the University of Michigan's Panel Study of Income Dynamics (PSID). A full account of the PSID, its history and main data files, can be found in Hill [22].

For our main sample, we have data on the earnings of fathers over the five-year period 1967–71, and the resources of sons are observed in 1991. We have 721 such parent and child pairs. The income concept taken here is the total labour income of the household head (measured in 1967 dollars – the year prior to which the survey was started). Fathers and sons are at least 25 years of age when their earnings are observed, and we have selected one child per parent family. The average age was 40.5 years for fathers in 1967, and 36.5 years for sons in 1991. Finally, as earnings are bound to vary over the life cycle in a non-random way, we have run prior regressions of the logarithm of labour income on age and age squared of the household head in each given year, and have chosen to work with the residuals from these initial regressions in the results reported below.

A first look at the covariance/correlation matrix of fathers' age-adjusted earnings is called for in order to have an idea as to what kind of error process may be postulated for the x data. The diagonal elements of Table 2A contain the cross-sectional variances. The above diagonal (s, τ) entries contain covariances, while the corresponding correlations between year s and year τ earnings are reported in the (τ, s) sub-diagonal entries. Though we shall not be dealing with this topic until Section 6, we note that there may be some non-stationarity in the data as the cross-sectional variance oscillates between 0.43 (1968) and 0.49 (1969). The one year auto-correlation is somewhere around 0.85 while the two-year correlation averages 0.80. Beyond two years, the pattern of decline is slower, with the 4-year correlation taking a value of 0.72. Because the drop in the auto-correlation is most rapid in the first two subsequent years, it is unlikely that an $AR(1)$ specification for the error component will adequately summarize the covariance structure of fathers' earnings. Conversely, because the auto-correlation declines in the ensuing years, an $MA(2)$ structure seems also too tight. In this sense, an $ARMA$ error process, nesting both AR and MA specifications may constitute a more flexible attempt for modelling the error process of earnings x .

Table 2A The correlation / covariance matrix of fathers' earnings in levels

| | 1967 | 1968 | 1969 | 1970 | 1971 |
|------|-------|-------|-------|-------|-------|
| 1967 | 0.447 | 0.376 | 0.381 | 0.350 | 0.328 |
| 1968 | 0.860 | 0.429 | 0.404 | 0.364 | 0.336 |
| 1969 | 0.809 | 0.876 | 0.496 | 0.394 | 0.367 |
| 1970 | 0.789 | 0.839 | 0.843 | 0.440 | 0.375 |
| 1971 | 0.718 | 0.752 | 0.764 | 0.828 | 0.466 |

The income concept is total labour income of the household head, measured in 1967 dollars.

Correlations appear as the sub-diagonal (s,t) entries pertaining to the (t,s) covariances.

Table 2B The correlation / covariance matrix of fathers' earnings in first differences

| | 1968 | 1969 | 1970 | 1971 |
|------|--------|--------|--------|--------|
| 1968 | 0.102 | -0.021 | -0.006 | -0.004 |
| 1969 | -0.205 | 0.104 | -0.058 | -0.004 |
| 1970 | -0.047 | -0.472 | 0.145 | -0.038 |
| 1971 | -0.030 | -0.029 | -0.258 | 0.152 |

The year t variable is $\Delta x_t = x_t - x_{t-1}$ where x_t are earnings at time t .

Correlations between Δx_s and Δx_t appear as the sub-diagonal (s,t) entries pertaining to the (s,t) covariances.

In Table 2B we report the covariance/correlation matrix of first time differences Δx_{is} . Once again we note that the variance of Δx_{is} increases over time. However, unlike the levels data, we notice that passed two periods, the correlation between Δx_{is} and Δx_{it} is very small. Thus, as a first approximation, a non-stationary $MA(1)$ process may provide an adequate modelling framework for the first difference in x . The fact that sub-diagonal elements of Table 2B are all negative suggests that the parameter θ_1 of the $MA(1)$ error process for Δx_{is} is negative.

Section 6 builds non-stationary models for Δx_{is} and for differences in differences of x . Because in practice such processes require more time-series variation in the data than processes estimated in levels, we have used a sub-sample of the Table 2A data with the additional sampling restriction that parents report non-zero earnings over the six year period 1967–1972. There are 694 such observations.

5 Evidence from the PSID

Our baseline regression is that of the child's 1991 age-adjusted earnings on those of the father in 1967. For this regression we initially consider two estimators: OLS and IV, which are intended to bracket the true value for the intergenerational elasticity β . In our sample these take on respectively the values 0.36 and 0.59, reported in the top line of Table 3. Next, we quantify the biases of these estimators by considering auto-regressive processes (Section 5.1), moving average processes (Section 5.2) and $ARMA$ processes (Section 5.3) for the error component u of the father's earnings x . In Section 5.4 we test for mis-specification of the IV estimator. Section 5.5 deals with tests of over-identifying moment restrictions.

5.1 $AR(1)$ process

Under the assumption that the transitory component of earnings follows an $AR(1)$ process, we use Eq. 22 for $\tau = 2, \dots, 5$ and $s = 1$ (where year 1 pertains to 1967 earnings) in order to estimate the variance of permanent income. That is, each of $(x_{i2} - \phi x_{i1})x_{i1}, \dots, (x_{i5} - \phi x_{i4})x_{i1}$ can be used in the numerator of Eq. 22, and an average of these is taken in order to estimate $E(\eta_p^2)$. We discuss in turn the estimates of the parameters $\phi, \lambda, b, \gamma, \rho, \delta_0$ and δ_1 reported in Table 3.

- The parameter ϕ of the auto-regressive component is estimated at 0.08, suggesting that serial correlation in errors of measurement is minor (a conclusion

Table 3 Estimation results

| | OLS | IV | λ | γ | ρ | ϕ | δ_0 | δ_1 | b |
|---------------|---------------|---------------|-----------|----------|--------|---------------|---------------|---------------|---------------|
| Cross-section | 0.356 (0.041) | 0.586 (0.072) | | | | | | | |
| AR(1) | | | 0.797 | 1.312 | 0.647 | 0.078 (0.073) | 0.346 (0.062) | 0.045 (0.017) | 0.447 (0.051) |
| MA(0) | | | 0.803 | 1.322 | 0.645 | | 0.342 (0.061) | 0.046 (0.017) | 0.444 (0.050) |
| MA(1) | | | 0.790 | 1.300 | 0.650 | | 0.352 (0.062) | 0.044 (0.017) | 0.451 (0.051) |
| MA(2) | | | 0.759 | 1.249 | 0.663 | | 0.378 (0.067) | 0.039 (0.018) | 0.470 (0.053) |
| ARMA(1,1) | | | 0.703 | 1.158 | 0.689 | 0.706 (0.321) | 0.435 (0.095) | 0.029 (0.022) | 0.506 (0.058) |
| ARMA(1,2) | | | 0.611 | 1.006 | 0.739 | 0.714 (0.431) | 0.579 (0.139) | 0.001 (0.031) | 0.583 (0.067) |

Standard errors appear inside parentheses.

that will shortly be reversed). With a standard error of 0.07, this estimate is not statistically different from zero.

- The parameter λ is then constructed as the estimate of $E(\eta_p^2)$ divided by the sample variance of 1967 earnings (see Eq. 23). In our data this produces an estimate of 0.80 for the signal to total variance ratio. This is suggestive of a moderate bias for the OLS estimator in comparison to conclusions reached by other authors in this respect. Though Bowles [11] reports estimates for λ ranging between 0.70 and 0.83 using various income concepts, Zimmerman [30] estimates this ratio to be 0.73 for wages and 0.66 for earnings, while Abul Naga [1] estimates this quantity to be in the range of 0.57 for family incomes and 0.62 for earnings.
- Our estimate of λ in turn can be used to correct the bias of the OLS estimator using Eq. 27. Given the correction factor underlying the $AR(1)$ model findings, the estimator b implies a value of 0.45 for the intergenerational elasticity β .
- Using Eq. 28, we evaluate the large sample bias of the instrumental variables estimator. The estimate of γ corresponding to the $AR(1)$ scenario is equal to 1.31, implying that $\hat{\beta}_{IV}$ over-estimates the intergenerational elasticity by 30%.
- The moment condition 22 providing an estimate of the variance of η_p , is used again in the estimation of the correlation ρ between education and permanent income (cf. Eq. 24). The estimate of ρ corresponding to the second line of Table 3 is equal to 0.65. We are not aware of other estimates of this parameter based on micro data. The simple correlation between education and age-adjusted annual earnings (i.e. $\lambda^{1/2}\rho$), taking a value of 0.58 in our 1967 data, would typically provide a lower bound estimate for ρ , since the presence of an earnings transitory variance component would inflate the denominator of such an expression.
- The parameters δ_0 and δ_1 of the structural model 10 for the determinants of the child's permanent income are estimated using Eqs. 25 and 26. As $\hat{\delta}_0$ and $\hat{\delta}_1$ are complicated functions of the sample moments $m(x_{is}x_{i\tau})$, $m(x_{is}e_i)$, etc., the *delta* method is used here in order to derive expressions for their standard errors (see Appendix 1 of the paper for further details). δ_0 is estimated at 0.35 with a standard error of 0.06 while δ_1 , the effect of the father's education on the child's permanent income, is estimated at 0.05 with a standard error of 0.02. It may be noted from this second finding that $\hat{\delta}_1$ is statistically different from zero – a point we shall take up again below.

5.2 Moving average processes

Given the near zero above estimate of ϕ , a first hand approximation to the $AR(1)$ solution is provided by the $MA(q)$ model estimates where $q = 0$, i.e. a model where errors of measurement are taken to be uncorrelated.

MA(0) process: uncorrelated errors

The $MA(0)$ process is the sub-case of Eq. 22 where $\phi = 0$ and $\tau > s$ (given $q = 0$). The results pertaining to this model are reported in the third line of Table 3. The variance of permanent income, $E(\eta_p^2)$, is now estimated as an average of the sample

covariances $m(x_{it}x_{is})$ for $\tau = 2, \dots, 5$. It may be noted that the solution is indeed very similar to the $AR(1)$ estimates. For instance, λ is estimated at 0.80 in both specifications, and the percentage bias of IV is only one point larger in the $MA(0)$ solution.

MA(1) process

In the next line of Table 3, the $MA(1)$ estimates, we allow the correlation between transitory earnings to be non-null in any two consecutive years, but we assume that it is zero for data observed two or more years apart. The $MA(1)$ estimates will therefore be consistent under the assumption that the $MA(0)$ specification is correct. The reverse however is not true, since the estimate of λ would then be inflated in the numerator by the covariance between transitory earnings components of years s and $s + 1$. The estimate of the variance ratio λ however drops only by 1%, to 0.79, the bias of IV is 30% and the remaining parameter estimates are essentially similar to those of the second and third lines of Table 3.

MA(2) process

This same exercise is further repeated in the fifth line of Table 3, under the assumption that errors of measurement follow an $MA(2)$ process. The estimate of λ now drops to 0.76 when the covariances between 1967 and 1970, and 1967 and 1971 earnings are used to evaluate the numerator of Eq. 23. As a result, b re-scales the 0.36 OLS estimate of β by a factor of 1.32 to arrive at an estimate of 0.47 for the intergenerational elasticity. The percentage bias implied by the instrumental variables estimator appears smaller, and γ is now estimated at 1.25, rather than in the earlier range of 1.30.

5.3 *ARMA* processes

Some changes are to be noted when we consider the more general class of *ARMA* error processes.

ARMA(1,1) process

Starting with the $ARMA(1, 1)$ results, the sixth line of Table 3, we note that ϕ , the parameter of the auto-regressive component of the error term, is estimated at 0.71 and is statistically significant (it was estimated at 0.08 under the $AR(1)$ assumption). In turn, λ is estimated at 0.70, bringing it more in line with other available estimates for the signal to total variance ratio. The estimate of ρ , the correlation between education and permanent income, is also revised upward from the earlier 0.65 range, to a value of 0.69. The bias of IV is now 16%, and the consistent adjusted least squares estimate of the intergenerational elasticity, b , takes a value of 0.51.

ARMA(1,2) process

Next, moving on to the $ARMA(1, 2)$ results, the estimate of λ drops further, to 0.61. Of more direct concern to the question addressed in this paper, the percentage bias of IV becomes virtually nil (1%) as a result of the downward revision of the

estimate of λ Eq. 23. In presence of $ARMA(1, 2)$ errors the consistent adjusted least squares estimator b estimates the intergenerational elasticity at 0.58. This finding is very much in line with the IV estimate of 0.59 (first line of Table 3). This last set of results therefore reverses the earlier conclusions of the $AR(1)$ and $MA(0)$ to $MA(2)$ specifications which pointed to something in the order of a 20% bias for OLS, and a 25% to 30% bias for IV (in the opposite direction).

5.4 Mis-specification of the IV estimator

Of particular relevance to our discussion is a test of the assumption that $\hat{\beta}_{IV}$ is inconsistent and biased away from zero. Inspecting Eq. 18, it may be noted that δ_1 equals zero either if $p \lim(\hat{\beta}_{IV}) = p \lim(\hat{\beta}_{OLS})/\lambda = \beta$, or if the correlation between education and permanent income is zero (a very unlikely assumption given many years of research on human capital). Assuming a positive correlation between education and permanent income, a rejection of the null hypothesis that $\delta_1 = 0$ would imply that the instrumental variables estimator is inconsistent. It may be inferred from the results of Table 3 that the t -statistic for δ_1 is above 2.0 for the $AR(1)$ and $MA(0)$ to $MA(2)$ results, in the order of 1.30 for the $ARMA(1, 1)$ solution and 0.03 for the $ARMA(1, 2)$ estimates.

If the alternative to the null is the assumption that $\delta_1 > 0$ – i.e. that $\hat{\beta}_{IV}$ is upwardly inconsistent, then the critical size of the test at the 5% level is 1.64 rather than 1.96. Evidence supporting the postulate that the instrumental variables estimator is biased away from zero is therefore only limited: it does not carry through when errors of measurement are modelled as $ARMA$ processes. We have more faith however in the $ARMA(1, q)$ test results since the related estimators are consistent either when the $AR(1)$ or $MA(q)$ assumptions are valid. The reverse, the $AR(1)$ or $MA(q)$ solutions being consistent when the error process is an $ARMA(1, q)$, is however not true. On the basis of our results therefore we wish to cast some doubt on the presumption that the IV estimator is upwardly inconsistent.

5.5 Tests of over-identifying restrictions

There is a last question we wish to address before we close this section of the paper. If the $ARMA(1, q)$ assumption is valid, this also means that $x_{it} - \phi x_{it-1}$ ($\tau > s + q$) provides a *within equation* instrument in a regression of y_{it} on x_{is} [19]. To distinguish this new estimator from $\hat{\beta}_{IV}$ (where education featured as an out of equation instrument) we shall denote the panel data IV estimator as $\hat{\beta}_{P-IV}$.

An important difference is also to be noted between panel data IV estimators and consistent adjusted least squares procedures such as Eq. 27. For the latter to be consistent, a full knowledge of the covariance structure of the error process is required. On the other hand, $\hat{\beta}_{P-IV}$ is consistent as long as the underlying instrument is uncorrelated with the error component u of observed income x .¹⁴

¹⁴For example, if the signal to total variance ratio λ varies with the age of parents, a consistent adjusted least square procedure which takes λ to be constant in the cross-section will not consistently estimate β ; a square Grawe [18] refers to as a life-cycle bias. The estimator $\hat{\beta}_{P-IV}$ however is robust to such a bias.

Table 4 Sargan tests for panel data IV estimators

| | $x_5 - \phi x_4$ | $x_4 - \phi x_3$ | $x_3 - \phi x_2$ | $x_2 - \phi x_1$ | x_4 | x_3 | x_2 | β | χ^2 [P -val] |
|-------------------|------------------|------------------|------------------|------------------|-------|-------|-------|---------------|-------------------------|
| <i>ARMA</i> (1,2) | + | + | | | | | | 0.536 (0.094) | 0.003 [0.957] |
| <i>ARMA</i> (1,1) | + | + | + | | | | | 0.488 (0.066) | 0.529 [0.768] |
| <i>MA</i> (2) | + | | | | + | | | 0.454 (0.051) | 0.450 [0.502] |
| <i>MA</i> (1) | + | | | | + | + | | 0.438 (0.049) | 1.319 [0.517] |
| <i>MA</i> (0) | + | | | | + | + | + | 0.425 (0.051) | 1.798 [0.615] |
| <i>AR</i> (1) | + | + | + | + | | | | 0.504 (0.066) | 2.880 [0.411] |

A + sign indicates that the variable is included in the set of instruments.

Standard errors appear inside parentheses. P -values appear inside square brackets.

The resulting panel data IV estimate when $x_{i5} - \phi x_{i4}$ and $x_{i4} - \phi x_{i3}$ ¹⁵ are used to instrument x_{i1} produces an estimate of 0.54, with a standard error of 0.09 (first line of Table 4). With $x_{i5} - \phi x_{i4}$ providing an additional instrument, the *ARMA*(1, 2) specification is testable by means of Sargan-type tests of over-identifying restrictions (Godfrey [16]; pp. 168–174). The resulting χ^2 test statistic (with a single degree of freedom) takes on a value of 0.003, corresponding to a P value of 0.96, suggesting that the *ARMA*(1, 2) assumption is very plausible. We may then be tempted to evaluate the plausibility of more restrictive assumptions pertaining to the error process. For instance, if Z_i is the set of instruments used to estimate β , in presence of *ARMA*(1, 1) errors Z_i is extended to include $x_{i3} - \phi x_{i2}$. The resulting panel IV estimator of β takes a value of 0.49, and the P value of the test is 0.77. Accordingly, the *ARMA*(1, 1) assumption is not rejected.

In the following line of Table 4 we test the *MA*(2) assumption. Partition Z_i into Z_{i1} and Z_{i2} , and let $Z_{i1} = x_{i5} - \phi x_{i4}$. While $Z_{i2} = [x_{i5} \ x_{i4}]$ provides two valid instruments in the *MA*(2) context, these cannot be jointly tested for exogeneity when $Z_{i1} = x_{i5} - \phi x_{i4}$, as the latter is a linear combination of the columns of Z_{i2} . Accordingly, in the Table we test for the exogeneity of x_{i4} . While the test does not reject the exogeneity assumption (its P value is in the order of 0.5), it is important to note that the panel data IV estimate of β drops from 0.54 (in the *ARMA*(1, 2) context) to 0.45. Recalling that an IV estimator that utilizes Z_{i1} and Z_{i2} can be decomposed into a weighted sum of the IV estimator based on Z_{i1} , and one based on the latter, this result shows that using future earnings (that is, in levels) as an instrument depresses considerably the panel data estimate of β . The over-identifying instruments pertaining to the *MA*(1), *MA*(0) and *AR*(1) assumptions are subsequently tested. The results are similar to those of the *MA*(2) tests: we do not reject the exogeneity assumption and the $\hat{\beta}_{P-IV}$ estimate of β is in the range of 0.40 to 0.50.

The failure to reject any of the assumptions tested in Table 4 is partly due to the large sampling variance of the $\hat{\beta}_{P-IV}$ estimate of the *ARMA*(1, 2) model. The following account may provide some intuition for understanding these test results: with an estimate of 0.54, and a standard error of 0.09, the panel data IV estimator of the *ARMA*(1, 2) specification spans a confidence interval for β with values ranging

¹⁵In the discussion that follows ϕ is replaced by its estimate of 0.714 obtained from the *ARMA*(1, 2) specification (cf. last line of Table 3).

between 0.35 and 0.72. The $MA(q)$ and $AR(1)$ panel data estimates reported in Table 4 are however all within the tighter range of 0.43 to 0.50, making it less likely for the test to detect departures from the exogeneity assumptions.

In a further attempt to discriminate between the various time series assumptions pertaining to the error term, we have considered an alternative test, motivated by the discussion contained in Arellano and Bond [6]. Note first that $\hat{\beta}_{P-IV}$ may be used as a preliminary estimator in the construction of more efficient GMM estimators. In the present panel data context, the efficient GMM estimator is essentially a weighted two-step IV method.¹⁶ A heteroscedasticity robust version of the Sargan test (see Eq. 10 of Arellano and Bond) has been considered, as evidence presented by the authors suggested that this form was more powerful in rejecting incorrect assumptions than the standard test of over-identifying restrictions. On the whole though, our data have produced very close GMM estimates, and the test results were qualitatively similar to those of Table 4.¹⁷

Though we were not able to reject the validity of the over-identifying restrictions related to simpler error structures, we may summarize our results by noting that they largely conform to the predictions of Proposition 3¹⁸: under the more general $ARMA(1, 2)$ specification the bias of OLS is considerably large, while that of $\hat{\beta}_{IV}$ is virtually zero. The bias of OLS drops, and that of IV increases in the $ARMA(1, 1)$ specification, and the pattern is further confirmed in the case of the nested simpler error structures. Furthermore, in the light of the specification test based on $\hat{\delta}_1$, evidence supporting the claim that education results in an upwardly biased IV estimator is only found in relation to the $AR(1)$ and $MA(0)$ to $MA(2)$ model estimates.

6 Some research directions

If we relax the assumption that permanent income is a time-invariant quantity, we can attempt to estimate the intergenerational elasticity under less restrictive assumptions. Specifically, introducing time variation in the permanent incomes of children and parents allows us to identify the structural model 10 in presence of correlated heterogeneity, that is to relax Eq. 11 (see Arellano [5], pp. 51–52 for a general discussion). We therefore introduce time subscripts for permanent incomes below, and we replace Eqs. 2 and 3 with the following assumptions

$$y_{it} = \eta_{cit} + v_{it} \tag{29}$$

$$x_{is} = \eta_{pis} + u_{is} \tag{30}$$

We also generalize the earnings process for x in several respects. Firstly, we shall introduce a stochastic process for permanent income, writing η_{pis} as the sum of last

¹⁶Squared residuals from the initial regression are used to provide the relevant weights.

¹⁷The P -values of the tests corresponding to those of the first to sixth line of Table 3 are respectively 0.956, 0.693, 0.470, 0.358, 0.426 and 0.428.

¹⁸In the context of our $ARMA(1, 2)$ error process, ϕ is estimated at 0.714 (Table 3, last line), while $\hat{\theta}_1 = -0.0035$, and $\hat{\theta}_2 = 0.2118$ (results not shown). Thus, we have $\hat{\phi}(\hat{\phi} + \hat{\theta}_1)\hat{\theta}_2 = 0.107$ and $\hat{\phi}[(\hat{\phi} + \hat{\theta}_1)\hat{\theta}_2 + \hat{\theta}_1] = 0.105$. These estimates thus verify the sufficient conditions of the Proposition.

period's permanent income and a disturbance. Secondly, we shall allow for non-stationarity in the earnings process, by assuming that the variance of the permanent income disturbance varies over time.

There are several insights to be obtained from such an exercise. Firstly, we show that in presence of correlated heterogeneity, the inequality 13 regarding the OLS and IV estimators and the parameter β need no longer hold (Section 6.1). Second, we show that in presence of correlated heterogeneity, the parameter β becomes somewhat uninformative about the structural relation between the permanent incomes of the parent and child families, even so when $\delta_1 = 0$. Finally, by introducing children specific permanent income trends (thus generalizing further the structural relation 10) we are able to explain why the OLS estimate 8 of the parameter β rises as children age (Section 6.3) without having to appeal to the Haider and Solon [28] hypothesis that measurement error is mean-reverting. Thus, the Section is also of independent interest in that it shows that a panel data modelling approach is rich enough to explain a major empirical puzzle of the literature within a classical errors-in-variables framework.

Noting that our estimate of ϕ in Table 3 for the general *ARMA* process is in the order of 0.7 with a standard error of 0.3 to 0.4, we shall begin this section with a simplifying assumption, namely that $\phi = 1$. Then, substituting this value for ϕ in Eq. 20, would amount to writing the first difference in x as a general *MA* process:

$$\Delta x_{is} = \varepsilon_{is} + \theta_1 \varepsilon_{is-1} + \cdots + \theta_q \varepsilon_{is-q} \quad (31)$$

Now that permanent income is allowed to vary over time, assume that it is the first difference in permanent income, rather than the first difference in earnings, which exhibits a general moving average structure:

$$\eta_{pis} = \eta_{pis-1} + \xi_{is} \quad (32)$$

$$\xi_{is} = \varpi_{is} + \kappa_1 \varpi_{is-1} + \cdots + \kappa_l \varpi_{is-l} \quad (33)$$

Together, Eqs. 30, 32, 33) would imply the following process for the first difference in x :

$$\Delta x_{is} = \varpi_{is} + \kappa_1 \varpi_{is-1} + \cdots + \kappa_l \varpi_{is-l} + \Delta u_{is} \quad (34)$$

Thus, if we abstract from measurement error ($\Delta u_{is} = 0$), and $l = q$, Eqs. 34 becomes a process observationally equivalent to Eqs. 31 or 20 when $\phi = 1$.

The present context is also useful to consider in that it completes the discussion of Section 3 in studying the identification of the structural model 10 under the assumption that $\phi = 1$; a case that could not be handled in the context of a model where the permanent incomes of parents and children are time-invariant quantities.

As shown in Appendix 3, in what follows it will be sufficient for identification purposes to model measurement error in x as an *MA*(q) process:

$$u_{is} = \varepsilon_{is} + \theta_1 \varepsilon_{is-1} + \cdots + \theta_q \varepsilon_{is-q} \quad (35)$$

where ε_{is} are *iid* error terms, uncorrelated with ϖ_{it} ,

$$E(\varepsilon_{it} \varepsilon_{is}) = 0 \quad s \neq t \quad (36)$$

$$\text{var}(\varepsilon_{is}) = \sigma_\varepsilon^2 > 0 \quad \text{for all } s \quad (37)$$

$$E(\varpi_{it} \varepsilon_{is}) = 0 \quad \text{for all } t, s \quad (38)$$

and to consider the permanent income process in the context of $l = 1$ (viz. an $ARMA(1, 1)$ process), where the permanent income shock ξ_{is} is non-stationary:

$$\xi_{is} = \varpi_{is} + \kappa_1 \varpi_{is-1} \tag{39}$$

$$E(\varpi_{is} \varpi_{it}) = 0 \quad s \neq t \tag{40}$$

$$E(\varpi_{is} \varpi_{it}) = E(\varpi_s^2) \quad s = t \tag{41}$$

The general stochastic process for earnings x we shall retain below therefore consists of Eqs. 30, 32, 35 and 39. Variances and covariances of disturbances are given by Eqs. 36–38 and Eqs. 40–41.

Modelling permanent income as an $ARMA(1, 1)$ process with a unit root ($\phi = 1$) will provide us with a platform to begin to explore the estimation of the structural model of income transmission using data in differences, as a means to control for various forms of unobserved heterogeneity.

6.1 Correlated random effects

As a first extension of the model of Section 2, we presently consider a structural model

$$\eta_{cit} = \delta_0 \eta_{pis} + \delta_1 e_i + v_i \tag{42}$$

with the assumptions

$$E(v_i e_i) \neq 0 \quad \text{and} \quad E(v_i \eta_{pis}) \neq 0 \tag{43}$$

In the above model the component v_i performs as a family specific random effect, which is posited to be correlated with the explanatory variables.

In a family utility model where parents are able to undertake efficient investments in their children’s human capital, where the inheritance of ability follows a first order Markov process, and permanent income is a time-invariant concept, the resulting relation between η_p and η_c is given by Eq. 1 with the important difference that $E(\zeta_i | \eta_{pi}) \neq 0$ [8]. Thus, returning to our specification 42, if $\delta_1 = 0$ (as in the Becker and Tomes efficiency model), we could equally envisage that $E(v_i | \eta_{pis}) < 0$ because v_i is a function of a luck term which together with parental ability determines η_{pis} . More generally δ_1 could be non-zero because parental investments are not efficient and, or, parental education is a useful index of family background, which has an impact on the child’s attainment. Also, the scenario $E(v_i e_i) \neq 0$ is easily conceivable: it arises for instance when v_i captures the effect of omitted determinants of η_{cis} , and such effects are correlated with education.¹⁹

In the present context, the parameters β , δ_0 and δ_1 become related through the following relation

$$\beta = \delta_0 + \frac{\delta_1 E(\eta_{ps} e) + E(\eta_{ps} v)}{E(\eta_{ps}^2)} \tag{44}$$

¹⁹Other sources of heterogeneity may pertain to the degree of parental altruism, and to whether parents are able to undertake efficient investments in their children’s human capital. The various sources of heterogeneity affect intergenerational earnings and consumption inheritance differently. See Han and Mulligan [21] for a discussion.

and the probability limits of the OLS and IV estimators are given by

$$plim(\hat{\beta}_{OLS}) = \frac{\delta_0 E(\eta_{ps}^2) + \delta_1 E(\eta_{ps}e) + E(\eta_{ps}v)}{E(\eta_{ps}^2) + E(u^2)} \quad (45)$$

$$plim(\hat{\beta}_{IV}) = \delta_0 + \frac{\delta_1 E(e^2) + E(ev)}{E(\eta_{ps}e)} \quad (46)$$

It is to be noted now that with the assumption Eq. 43 replacing Eq. 11, the inequality 13 bounding the parameter β is no longer valid: while the OLS estimator remains biased toward zero, it can no longer be claimed that the IV estimator is biased upwards;²⁰

Proposition 4 *Let the structural model mapping the permanent incomes of parents and children be given by Eqs. 42–43. Then the inequality Eq. 13 need no longer hold.*

Note furthermore that under the present set-up, it is no longer the case under the null hypothesis $\delta_1 = 0$ that β coincides with δ_0 . While δ_0 remains a slope parameter, it is no longer straightforwardly clear what interpretation ought to be given to β in the light of Eq. 44, and what is its exact relevance as an index of justice attainment. As such, our interest ought to be recentered around a direct estimation of δ_0 and δ_1 from the measurement model pertaining to Eq. 42.

Replacing Eqs. 29 and 30 into Eq. 42 we obtain the required relation:

$$y_{it} = \delta_0 x_{is} + \delta_1 e_i + v_i + v_{it} - \delta_0 u_{is} \quad (47)$$

If we were to take first differences in Eq. 47 we would, eliminate v_i to obtain the estimator

$$\hat{\delta}_{0,FD} = \frac{\sum_i (y_{it} - y_{it-1})(x_{is} - x_{is-1})}{\sum_i (x_{is} - x_{is-1})^2} \quad (48)$$

which is purged from heterogeneity biases. To estimate δ_1 one would need to construct a further regression. Define $y_i^+ = \bar{y}_i - \hat{\delta}_{0,FD} \bar{x}_i$ and consider the regression

$$y_i^+ = \delta_1 e_i + v_i + \bar{v}_i - \delta_0 \bar{u}_i \quad (49)$$

Below we calculate an estimator $\hat{\delta}_{1,FD}$ obtained by fitting the above regression via OLS:

$$\hat{\delta}_{1,FD} = \frac{\sum_i y_i^+ e_i}{\sum_i e_i^2} \quad (50)$$

The estimators $\hat{\delta}_{0,FD}$ and $\hat{\delta}_{1,FD}$ remain inconsistent because of measurement error in x . Furthermore, because under Eq. 43, e_i is correlated with v_i , $\hat{\delta}_{1,FD}$ is inconsistent independently of measurement error biases. However, an instrumental variables

²⁰For instance if $\delta_1 = 0$ and

$$E(ev) \leq \frac{E(\eta_{ps}e)E(\eta_{ps}v)}{E(\eta_{ps}^2)}$$

then it is the case that $p \lim(\hat{\beta}_{IV}) \leq \beta$.

estimator using leads of x as instruments consistently estimates δ_0 (see Appendix 3). In particular, if measurement error u is an $MA(q)$ process, we consider the estimator

$$\tilde{\delta}_0 = \frac{\sum_i (y_{it} - y_{it-1}) x_{it}}{\sum_i (x_{is} - x_{is-1}) x_{it}} \quad \tau > s + q \tag{51}$$

to identify δ_0 in the present context.

By analogy with Eq. 49, δ_1 may be estimated by considering a further regression. Define then $\tilde{y}_i = \bar{y}_i - \tilde{\delta}_0 \bar{x}_i$ and consider the regression

$$\tilde{y}_i = \delta_1 e_i + v_i + \bar{v}_i - \delta_0 \bar{u}_i \tag{52}$$

Since $E(v_i e_i) \neq 0$, if z_i is some variable correlated with e_i but uncorrelated with v_i and the remaining error terms, the estimator

$$\tilde{\delta}_1 = \frac{\sum_i \tilde{y}_i z_i}{\sum_i e_i z_i} \tag{53}$$

will identify δ_1 .

6.2 Individual specific trends

There may also be some interest in generalizing Eq. 42 further in order to introduce an individual specific trend in the child’s permanent income:

$$\eta_{cit} = v_{0i} + v_{1i}t + \delta_0 \eta_{pis} + \delta_1 e_i \tag{54}$$

where v_{0i} is the intercept, and v_{1i} the slope, of the individual permanent income trend. We consider estimation of δ_0 and δ_1 under the assumptions

$$E(v_{ji} e_i) \neq 0 \quad \text{and} \quad E(v_{ji} \eta_{pis}) \neq 0 \quad j = 0, 1 \tag{55}$$

We maintain below that the measurement error component of parental earnings, u_{is} , is uncorrelated with the intercept and slope of the individual-specific permanent income trend:

$$E(v_{0i} u_{is}) = 0 \quad E(v_{1i} u_{is}) = 0 \quad \text{for all } s \tag{56}$$

The parameters β , δ_0 and δ_1 now become related through the following relation:

$$\beta = \delta_0 + \frac{\delta_1 E(\eta_{ps} e) + E(\eta_{ps} v_0) + t.E(\eta_{ps} v_1)}{E(\eta_{ps}^2)} \tag{57}$$

Under Eq. 54 both the intercept v_{0i} and slope v_{1i} of the time trend become individual specific. The main reason for considering such a specification is motivated by the findings of a recent literature [18, 24] which has widely documented the pattern that estimates of the intergenerational elasticity parameter β of the simple Galtonian model 1 rise as children’s earnings are observed at later stages of the life cycle.

If children raised in wealthier families are the ones who study longer, such persons typically begin their careers with low salaries and experience higher earnings growth over time. Under such circumstances we could conceive that $E(v_{0i} \eta_{pis}) \leq 0$ and $E(v_{1i} \eta_{pis}) \geq 0$. In addition, if wealthier parents are also highly educated, we could also envisage that $E(v_{0i} e_i) \leq 0$ and $E(v_{1i} e_i) \geq 0$.

If v_{it} is positively correlated with η_{pis} , and e_i then both the OLS and IV estimates of β would rise as children age:

$$plim(\hat{\beta}_{OLS}) = \frac{\delta_0 E(\eta_{ps}^2) + \delta_1 E(\eta_{ps}e) + E(\eta_{ps}v_0) + t.E(\eta_{ps}v_1)}{E(\eta_{ps}^2) + E(u^2)} \quad (58)$$

$$plim(\hat{\beta}_{IV}) = \delta_0 + \frac{\delta_1 E(e^2) + E(ev_0) + t.E(ev_1)}{E(\eta_{ps}e)} \quad (59)$$

Proposition 5 *Let the structural model mapping the permanent incomes of parents and children be given by Eqs. 54–55. Then:*

- (1) β rises as children age if $E(\eta_{ps}v_1) \geq 0$.
Furthermore, under the classical errors in variables assumptions 29–30,
- (2) $plim(\hat{\beta}_{OLS})$ rises as children age if $E(\eta_{ps}v_1) \geq 0$
- (3) $plim(\hat{\beta}_{IV})$ rises as children age if $E(ev_1) \geq 0$.

The above result is of direct relevance to the literature since it shows that, under the classical errors in variables assumptions, a rise in the $\hat{\beta}_{OLS}$ estimate as children age can be explained by the presence of a child specific permanent income trend. As such, while mean-reverting measurement error accounts well for this finding, it is not the only possible explanation.

Next consider the estimation of Eq. 54. The underlying measurement model takes the form:

$$y_{it} = v_{0i} + v_{1i}t + \delta_0 x_{is} + \delta_1 e_i + v_{it} - \delta_0 u_{is} \quad (60)$$

Taking a *difference in first differences* in the above equation eliminates the random parameters of the time trend:

$$\begin{aligned} \Delta y_{it} - \Delta y_{it-1} &= \delta_0 (\Delta x_{is} - \Delta x_{is-1}) \\ &\quad + (\Delta v_{it} - \Delta v_{it-1}) - \delta_0 (\Delta u_{is} - \Delta u_{is-1}) \end{aligned} \quad (61)$$

OLS estimators for the difference in first difference equations are given by $\hat{\delta}_{0,DFD}$ (OLS fitted to Eq. 61) and $\hat{\delta}_{1,DFD}$ obtained by fitting the equation

$$y_i^{++} = \delta_1 e_i + v_i + \bar{v}_i - \delta_0 \bar{u}_i \quad (62)$$

by ordinary least squares, where $y_i^{++} = \bar{y}_i - \hat{\delta}_{0,DFD} \bar{x}_i$.

Because of the problem of measurement error, consistent estimation of δ_0 and δ_1 further requires instrumentation of Eq. 61. Accordingly, when u_{is} is an $MA(q)$ process, the estimator

$$\tilde{\delta}_0 = \frac{\sum_i (\Delta y_{it} - \Delta y_{it-1}) x_{it}}{\sum_i (\Delta x_{is} - \Delta x_{is-1}) x_{it}} \quad \tau > s + q \quad (63)$$

is a consistent differences in first differences instrumental variables (DFD-IV) estimator (see Appendix 3 of the paper).

If we define the fitted slope

$$\hat{v}_{li} = \frac{1}{2} \left[(\Delta y_{it} + \Delta y_{it-1}) - \tilde{\delta}_0 (\Delta x_{is} + \Delta x_{is-1}) \right] \quad (64)$$

and

$$\tilde{y}_i = \bar{y}_i - \tilde{\delta}_0 \bar{x}_i - \hat{v}_{1i}t$$

then, the analogue of Eq. 52 is a regression

$$\tilde{y}_i = \delta_1 e_i + v_{0i} + \bar{v}_i - \delta_0 \bar{u}_i \tag{65}$$

which identifies δ_1 using some instrument z_i for education:

$$\tilde{\delta}_1 = \frac{\sum_i \tilde{y}_i z_i}{\sum_i e_i z_i} \tag{66}$$

6.3 Further results

We reconsider the estimation of the structural model of income transmission in presence of unobserved heterogeneity. As discussed in the above sub-sections, our efforts are now centered around a direct estimation of δ_0 and δ_1 . One new concern here is a preliminary evaluation of how well estimators based on differencing and difference in first differences perform in practice in identifying the above model.

In Table 5 we consider six estimators for the structural model of income transmission.

- The OLS estimator in levels abstracts from measurement error in x as well as potential correlated heterogeneity. Fitting Eq. 47 by OLS we obtain an estimate of 0.25 for δ_0 . This estimate is considerably smaller than the 0.36 OLS estimate of β in Table 3, as the coefficient on parental education is estimated to be significant (a 0.07 value with a standard error of 0.017).

Table 5 Levels, first difference and difference in first differences estimations

| | δ_0 | δ_1 | MA(1) measurement error | | MA(2) measurement error | |
|------------|-------------------|------------------|----------------------------|------------------|----------------------------|------------------|
| | | | δ_0 | δ_1 | δ_0 | δ_1 |
| OLS-levels | 0.254 (0.050) | 0.072 (0.017) | | | | |
| OLS-FD | 0.063 (0.050) | 0.103 (0.012) | | | | |
| OLS-DFD | -0.075 (0.061) | 0.119 (0.020) | | | | |
| IV-levels | | | 0.356 (0.065) | 0.053 (0.018) | 0.363 (0.071) | 0.052 (0.019) |
| IV-FD | | | -0.493 (0.629) | 0.373 (0.040) | -0.713 (0.817) | 0.454 (0.045) |
| IV-DFD | | | 0.347 (0.647) | 0.096 (0.052) | 0.385 (0.652) | 0.080 (0.052) |

IV-levels estimates are robust to measurement error but not to heterogeneity biases.

IV-FD estimates are robust to measurement error and correlated random effects.

IV-DFD estimates are robust to measurement error and child specific income trends.

- The next line of the Table, OLS-FD, reports the estimate of δ_0 obtained via Eq. 48. This estimator eliminates heterogeneity biases when it is taken that v is correlated with e and η_{ps} . Because first differencing aggravates measurement error problems, the resulting estimator is likely to be biased downwards to a greater extent than OLS in levels. The resulting estimate of 0.06, with a standard error of 0.05 however is hard to reconcile with a story of measurement error bias.
- The OLS-DFD estimates of δ_0 and δ_1 (cf. Eqs. 61 and 62) are robust to correlated heterogeneity in intercept and slope of the permanent income trend, but not to measurement error. The estimate of δ_0 (-0.08 with a standard error of 0.061) is again problematic.

It is thus necessary to reconsider the levels and differencing procedures in the light of measurement error. This is the purpose of the next three estimators which instrument the levels, first difference and difference in differences equations. There are two sets of estimates reported in each line of Table 5, depending on whether the error component u_{is} is modelled as an $MA(1)$ or $MA(2)$ process.

- The IV-levels estimator instruments 1967 earnings in Eq. 47 using available leads x_τ , where $\tau \geq s + 1$ in the $MA(1)$ specification, and $\tau \geq s + 2$ in the $MA(2)$ specification. The estimator is robust to measurement error, but not to the forms of correlated heterogeneity discussed in the sub-sections above. The estimates of δ_0 now rise from 0.25 for OLS in levels to 0.36 in the fourth line of results in Table 5.
- The IV-FD estimator is an instrumented version of OLS-FD. The estimates of δ_0 (-0.50 when u_s is modelled as an $MA(1)$ specification, and -0.71 in the $MA(2)$ specification) are again indicative of a specification bias.
- Finally, the IV-DFD estimator estimates δ_0 at 0.35 in the $MA(1)$ specification and 0.39 in the $MA(2)$ specification.

In this sense, the specification of a permanent income trend, allowing for individual specific intercepts and slopes appears in the light of these results to offer a more promising avenue for future research on the estimation of the structural model of income transmission in the context of correlated heterogeneity.

Note however one major drawback of differencing estimators in comparison to equations estimated in levels. The standard error of the estimate of δ_0 is large, thus resulting in a statistically not different from zero estimate. The result points to the fact that the time-series variation in earnings panels such as the PSID may turn out to be insufficient for the use of differencing procedures to be statistically informative for parent and child permanent income elasticities. However, this problem may be overcome in the context of other data sources where parent and child samples sizes are well over 700.

The estimates of δ_1 are obtained using auxiliary regressions based on Hausman and Taylor procedures. For the IV-FD (IV-DFD) estimator to be robust to the presence of correlated random effects (individual specific trends), the regressions 52 and 65 must be instrumented. To the extent that a race dummy is a correlate of education, but is orthogonal to the heterogeneity components, the resulting IV-FD and IV-DFD estimators of δ_1 will be consistent. If we set aside the IV-FD solutions, all estimates vary in the range $[0.05-0.12]$. The estimate of δ_1 is statistically significant in the OLS, levels, first difference and difference in first differences procedures, which do not treat the measurement error problem.

The IV-levels estimator of δ_1 is also significant. The IV-DFD estimate of δ_1 however has a t -statistic of 1.85 in the context of the $MA(1)$ specification and a t -statistic of 1.54 in the $MA(2)$ specification. This last result is certainly to be treated with caution. Nonetheless it points again to the fact that, having controlled for parental income, supporting evidence in favour of the claim that education is a determinant of the child's permanent income is fairly limited.

7 Conclusions

By providing repeated measurements on error-ridden variables, panel data allow the researcher to identify a wide range of errors in variables models, of which the Galtonian model of the intergenerational mobility literature is an example. We have used this result in the present paper in order to quantify the large sample biases of the ordinary least squares and instrumental variables estimators of the intergenerational elasticity. The key quantities in evaluating the biases of these estimators were shown to be the signal to total variance ratio and the correlation between education and permanent income, for which consistent estimators have been provided. As a by-product of our discussion we have proposed a test to investigate the null hypothesis that the instrumental variables estimator is consistent, against the alternative that it is inconsistent and biased away from zero.

In the paper we have estimated models of income transmission using panel data where earnings were measured in levels, in first differences and as differences in first differences. For the earnings data measured in levels, we have found that a tight specification of the auto-correlation in the error process (for instance $AR(1)$ and $MA(1)$ processes) tends to inflate the signal to total variance ratio, thus making the relative bias of the OLS estimator appear moderate, and that of IV large. The results of this paper suggest that the estimate of the signal to total variance ratio falls when we introduce more flexibility in the specification of the error process.

Estimators based on first differences in earnings are robust to biases arising from correlated random effects, but are sensitive to measurement error. In principle, if the estimate of the intergenerational elasticity is higher than in the levels data, this suggests that parental permanent income is negatively correlated with the unobserved random effect. Our estimates based on first differences, with and without correction for measurement error, turned out however to be unsatisfactory.

If the OLS and IV estimates increase as children age, this may be indicative of mean reverting measurement error, as has been suggested in several recent studies. An alternative explanation we have explored in Section 6, is that the rise in the estimates may be due to the presence of a child specific permanent income trend in the structural model of income transmission; for which difference in first differences estimation is a suitable empirical strategy. Estimates of the structural model of income transmission based on this latter approach were found to be more plausible than for data constructed in first differences, yet they exhibited large standard errors.

Acknowledgements I wish to thank Christophe Kolodziejczyk, Jaya Krishnakumar, Dirk Van de gaer and Tom Wansbeek for discussions. The comments of Mario Jametti and three referees are gratefully acknowledged. I am responsible for any errors or omissions.

References

1. Abul Naga, R.: Estimating the intergenerational elasticity of incomes: an errors in variables framework. *Economica* **69**, 69–91 (2002)
2. Abul Naga, R., Krishnakumar, J.: Panel data estimation of the intergenerational elasticity of incomes. In: Krishnakumar, J., Ronchetti, E. (eds.) *Panel Data Econometrics: Future Directions, Papers in Honour of Professor Pietro Balestra*. North-Holland, Amsterdam (2000)
3. Aigner, D., Hsiao, C., Kapteyn, A., Wansbeek, T.: Latent variable models in econometrics. In: Griliches, Z., Intriligator, M. (eds.) *Handbook of Econometrics*, vol. 2. Elsevier, Amsterdam (1984)
4. Altonji, J., Dunn, K.: Relationships among the family incomes and labor market outcomes of relatives. *Res. Labor Econ.* **12**, 269–310 (1991)
5. Arellano, M.: *Panel Data Econometrics*. Oxford University Press, Oxford (2003)
6. Arellano, M., Bond S.: Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations. *Rev. Econ. Stud.* **58**, 77–297 (1991)
7. Atkinson, A., Maynard, A., Trinder, C.: *Parents and Children: Incomes in Two Generations*. Heinemann, London (1983)
8. Becker, G., Tomes, N.: Human capital and the rise and fall of families. *J. Labor Econ.* **4**, S1–S39 (1986)
9. Behrman, J., Taubman, P.: The intergenerational elasticity between children's adult earnings and their parents' incomes: results from the Michigan panel study of income dynamics. *Rev. Income Wealth* **36**, 115–127 (1990)
10. Björklund, A., Jantti, M.: Intergenerational income mobility in Sweden compared to the United States. *Am. Econ. Rev.* **87**, 1009–1018 (1997)
11. Bowles, S.: Schooling and inequality from generation to generation. *J. Political Econ.* **80**, S219–S251 (1972)
12. Bowles, S., Gintis H.: The inheritance of inequality. *J. Econ. Perspectives* **16**, 3–30 (2002)
13. Dearden, L., Machin, S., Reed, H.: Intergenerational mobility in Britain. *Econ. J.* **107**, 47–66 (1997)
14. Ermisch, J., Francesconi, M.: Intergenerational mobility in Britain: new evidence from the British household panel survey. In: Corak M. (ed.) *Generational Income Mobility in North America and Europe*. Cambridge University Press, Cambridge (2004)
15. Fortin, N., Lefebvre, S.: Intergenerational income mobility in Canada. In: Corak M. (ed.) *Labor Markets, Social Institutions and the Future of Canada's Children*, Ottawa, Statistics Canada (1998)
16. Godfrey, L.: *Misspecification Tests in Econometrics*. Cambridge University Press, Cambridge (1988)
17. Goldberger, A.: Economic versus mechanical models of income transmission. *Am. Econ. Rev.* (1989)
18. Grawe, N.: Life cycle bias in the estimation of intergenerational earnings persistence. *Labour Econ.* **13**, 551–570 (2006)
19. Griliches, Z., Hausman, J.: Errors in variables in panel data. *J. Econometrics* **31**, 93–118 (1986)
20. Haider, S., Solon, G.: Life-cycle variation in the association between current and lifetime earnings. *Am. Econ. Rev.* **96**, 1308–1320 (2006)
21. Han, S., Mulligan, C.: Human capital, heterogeneity and estimated degrees of intergenerational mobility. *Econ. J.* **111**, 207–243 (2001)
22. Hill, M.: *The Panel Study of Income Dynamics: a User's Guide*. Sage Publications, New York (1993)
23. Hsiao, C., Taylor, G.: Some remarks on measurement error and the identification of panel data models. *Statistica Neerlandica* **45**, 187–194 (1991)
24. Mazumder, B.: Fortunate sons: new estimates of intergenerational mobility in the United States using social security earnings data. *Rev. Econ. Stat.* **87**, 235–255 (2005)
25. Meijer, E., Wansbeek T.: Measurement error in a single regressor. *Econ. Lett.* **69**, 277–284 (2000)
26. Mulligan, C.: Galton versus Human capital approach to inheritance. *J. Political Econ.* **107**, S184–S224 (1999)
27. Oreopoulos, P.: The long-run consequences of living in a poor neighborhood. *Quart. J. Econ.* **118**, 1533–1575 (2003)

28. Solon, G.: Intergenerational income mobility in the United States. *Am. Econ. Rev.* **82**, 393–408 (1992)
29. Wansbeek, T.: GMM estimation in panel data with measurement error. *J. Econometrics* **104**, 259–268 (2001)
30. Zimmerman, D.: Regression toward mediocrity in economic stature. *Am. Econ. Rev.* **82**, 409–429 (1992)