A Note on the Estimation of Intergenerational Income Correlations by the Method of Averaging

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

July 17, 2000

Abstract

Averaging methods are routinely used in order to limit biases resulting from the mismeasurement of permanent incomes. The Solon/Zimmerman estimator regresses a single-year measurement of the child's resources on a T-period average of the parents' income while the Behrman/Taubman estimator regresses an S-period average of the child's resources on a T-period average of the parents' income. The latter estimator is shown to be the arithmetic mean of the S slope estimates arising from the Solon/Zimmerman methodology. However, because sampling variation produces yearly changes in the variance of children's incomes, it is shown that the Behrman/Taubman estimator is not efficient in the class of estimators which can be expressed as a weighted sum of the S distinct Solon/Zimmerman estimates. The minimum variance estimator in the above class is thus derived and applied to a US sample.

Keywords: Intergenerational mobility, measurement error, averaging methods, minimum variance estimation.

JEL codes: I3, J62.

*I am grateful to an anonymous referee who prompted the question addressed in this paper. I also wish to thank Gerard Antille for discussions. 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

It is a well documented fact that estimates of the intergenerational correlation of incomes based on single year measurements on the incomes of parents and children are biased towards zero. Panel data however usually provide several measurements on the incomes of parents and children which, when averaged over time, can substantially reduces biases resulting from measurement error. Two variants of the method of averaging are routinely used in the literature. The first of these consists in regressing a single year measurement of the child's income on a T-period average of the parent family's resources. Applications of this method can be found in Solon (1992), Zimmerman (1992) and others. The second variant, which first appeared in the work of Behrman and Taubman (1990), regresses an S-period average of the child's income on a T-period average of the parent family's resources. Recent applications of this second method may be found for instance in Mulligan (2000).

It is understood that both estimators will have the same probability limit as errors-in-variables biases occur as a result of mismeasurement of the explanatory variable, viz. parental income, but not because the dependent variable is also subject to measurement error. It is also expected that the second estimator, the Behrman and Taubman variant, be more efficient since it utilizes more information than the former. Beyond this however it is not known what analytical relation exists between these two estimators. Establishing the relation between the two estimators is useful for two separate reasons. On the one hand it allows us to interpret more clearly the different findings available from the two variants of the method of averaging. Secondly, as will be shown below, deriving this relation will allow us to define a general class of estimators for the intergenerational correlation based on the method of averaging, of which the Behrman and Taubman estimator is a member, yet it is not the most efficient estimator.

Below, we show that the Behrman and Taubman estimator is the arithmetic mean of the S slope estimates derived from the regression of the child's year-s income on the T-period average of the parent family's income. That is, it is a member of the class of estimators which can be expressed as a weighted sum of the S slope estimates derived from the Solon/Zimmerman variant of the method of averaging (with the particular feature that all weights are restricted to be constant and equal to 1/S). However, because sampling variation will produce year to year changes in the variance of children's incomes, it is unlikely that each of these S slope estimates will have equal sampling variance. Thus it is generally possible to derive a more efficient estimator than the Behrman and Taubman method, which attributes unequal weights to each of the S slope estimates derived from the Solon/Zimmerman variant of the method of averaging.

The plan of the note is the following. In section 2 we introduce our notation and estimators for the Galtonian model of income transmission. In section 3 we establish the arithmetic relation between the two variants of the method of averaging. There, we also derive the minimum variance estimator in the class of weighted sums of the Solon/Zimmerman estimates. Section 4 contains an illustrative example based on a US sample of parents and children extracted from the Panel Study of Income Dynamics. Section 5 concludes with a summary of the main points.

2. Framework and definitions

We are considering a relation where the researcher seeks to estimate a Galtonian model

$$\eta_{ci} = \beta\eta_{pi} + u_i \tag{1}$$

where η_{pi} and η_{ci} $(i = 1; \ldots; n)$ are respectively the logarithms of the permanent incomes of parents and children measured on deviation from their means, and n denotes sample size. The parameter β is the elasticity of the child's income with respect to that of her parents' resources, and here will be referred to as the intergenerational correlation.

It is assumed that η_{ci} and η_{pi} are unobserved. However, in a panel data environment the researcher will usually possess several noisy measurements y_{is} $(s = 1; \ldots; S)$ and x_{it} $(t = 1; \ldots; T)$ on the permanent incomes of parents and children

$$y_{is} = \eta_{ci} + \phi_{is} \tag{2}$$

$$x_{it} = \eta_{pi} + \varepsilon_{it} \tag{3}$$

such that ϕ_{is} and ε_{it} obey the classical errors in variables properties $E(\phi_{is}\eta_{ci}) = E(\phi_{is}\eta_{pi}) = E(\phi_{is}\varepsilon_{it}) = E(\varepsilon_{it}\eta_{pi}) = E(\varepsilon_{it}\eta_{ci}) = 0$: Furthermore, let σ_{pp} and $\sigma_{\varepsilon\varepsilon}$ denote the variances of the permanent and transitory income components pertaining to x_{it}: The probability limit of the ordinary least squares estimator of a regression of y_{is} on x_{it}, $\hat{\beta}_{st}$, takes the form

$$\operatorname{plim}(\hat{\beta}_{st}) = \beta\frac{\sigma_{pp}}{\sigma_{pp} + \sigma_{\varepsilon\varepsilon}} \tag{4}$$

which will hold for all t and s provided the variance of ε_{it} is time-invariant. Now define $\bar{x}_i = \sum_{t=1}^{T} x_{it}/T$ as a time-series average of individual observations on the parent family's income. A regression of y_{is} on $\bar{x}_i$ (the Solon/Zimmerman variant of the method of averaging) produces an estimator $\hat{\beta}_s$ with probability limit

$$\text{plim}(\hat{\beta}_s) = \beta \frac{\sigma_{pp}}{\sigma_{pp} + \sigma_{\varepsilon\varepsilon}/T} \tag{5}$$

so that the asymptotic bias of the resulting estimator diminishes as T increases. A second variant of the method of averaging, first used by Behrman and Taubman (1990) regresses a time-series average $\bar{y}_i = \sum_{t=1}^{T} y_{it}/T$ on $\bar{x}_i$. Let $\hat{\beta}_{ave}$ denote this estimator. Furthermore define $a_i = \bar{x}_i / \sum_i \bar{x}_i^2$. We may then express $\hat{\beta}_{ave}$ and $\hat{\beta}_s$ in the following compact notation:

$$\hat{\beta}_{ave} = \sum_i a_i \bar{y}_i \tag{6}$$

$$\hat{\beta}_s = \sum_i a_i y_{is} \tag{7}$$

Solon (1992), Zimmerman (1992), Björklund and Jantti (1997) and others use the estimator $\hat{\beta}_s$ while Behrman and Taubman (1990) and Mulligan (2000) use the form $\hat{\beta}_{ave}$. One natural question to ask is what statistical relation (if any) exists between $\hat{\beta}_{ave}$ and $\hat{\beta}_t$. We turn to this question in section 3 below.

3. A minimum variance averaging estimator

Replacing for $\bar{y}_i$ in (6), we obtain

$$\hat{\beta}_{ave} = \sum_i a_i (y_{i1} + \dots + y_{iS})/S = \frac{1}{S} \sum_{s=1}^{S} \hat{\beta}_s \tag{8}$$

That is, the estimator $\hat{\beta}_{ave}$ is the arithmetic mean of the various estimates $\hat{\beta}_s$. In particular, its probability limit will be identical to that of $\hat{\beta}_s$. However, since the latter variant of the method of averaging utilizes more information than the former, it is expected that $\hat{\beta}_{ave}$ will be a more efficient estimator than $\hat{\beta}_s$. Now consider the class of estimators

$$\hat{\beta}_w = \left\{ \sum_{s=1}^{S} w_s \hat{\beta}_s \;\middle|\; \sum_{s=1}^{S} w_s = 1 \right\} \tag{9}$$

$\hat{\beta}_w$ is a general form for the family of weighted estimators of the various $\hat{\beta}_s$ with probability limit equal to that of $\hat{\beta}_s$. In particular, $\hat{\beta}_{ave}$ is a member of the class $\hat{\beta}_w$ with constant and equal weights $w_s = \frac{1}{S}$ for all s.

Even if we are to assume that in the population y_{is} and $y_{is'}$ have identical variances, it is unlikely in the sample that the estimators $\hat{\beta}_s$ will have equal variances. Thus, while in general $\hat{\beta}_{ave}$ will be more efficient than a given $\hat{\beta}_s$, it is possible to consider more efficient variants of $\hat{\beta}_w$, which do not assign equal weights to each $\hat{\beta}_s$. In particular, the minimum variance estimator of the class $\hat{\beta}_w$ is solution to the problem

$$\min_{w_1,\ldots,w_S} \operatorname{var}(w_1\hat{\beta}_1 + \cdots + w_S\hat{\beta}_S) \tag{10}$$

subject to the constraint

$$\sum_{s=1}^{S} w_s = 1 \tag{11}$$

Define $\hat{\Omega}$ as the $S \times S$ sample covariance matrix of the $S \times 1$ column vector $\hat{\beta} = [\hat{\beta}_1 \cdots \hat{\beta}_S]'$. Let w be an $S \times 1$ column vector of weights and write the Lagrangean of the problem as follows:

$$L(w;\mu) = w'\hat{\Omega}w - 2\mu(w'\iota - 1) \tag{12}$$

where μ is the Lagrange multiplier associated with the constraint (11) and ι is a vector of ones.

Taking first order conditions in (12) we obtain

$$\frac{\partial L}{\partial w} = \hat{\Omega}w - \mu\iota = 0 \tag{13}$$

$$\frac{\partial L}{\partial \mu} = w'\iota - 1 = 0 \tag{14}$$

Pre-multiplying (13) by w' we obtain $\mu = w'\hat{\Omega}w$, which upon solving gives us the weights $w = (\iota'\hat{\Omega}^{-1}\iota)^{-1}\hat{\Omega}^{-1}\iota$. The minimum variance estimator in the class (9) therefore takes the generalized least squares form

$$\hat{\beta}^*_w = (\iota'\hat{\Omega}^{-1}\iota)^{-1}\iota'\hat{\Omega}^{-1}\hat{\beta} \tag{15}$$

It is instructive to consider the case $S = 2$, the situation in which the researcher possesses two measurements on the child's income. This gives a well known formula (Fraser, 1976; ch. 9, section C)

$$\hat{\beta}^*_w = \frac{\hat{\sigma}_{22} - \hat{\sigma}_{12}}{\hat{\sigma}_{11} + \hat{\sigma}_{22} - 2\hat{\sigma}_{12}}\hat{\beta}_1 + \frac{\hat{\sigma}_{11} - \hat{\sigma}_{12}}{\hat{\sigma}_{11} + \hat{\sigma}_{22} - 2\hat{\sigma}_{12}}\hat{\beta}_2 \tag{16}$$

with the property that the larger the sample variance $\hat{\omega}_{22}$ of $\hat{\delta}_2$ is, the more weight $\hat{\delta}_1$ will be given in the calculation of the minimum variance estimator β^*_w. Equal weighting is only optimal in the case where $\hat{\omega}_{11} = \hat{\omega}_{22}$.

In general then, $\hat{\beta}_{ave}$ is dominated by β^*_w in the sense that

$$\mathrm{var}(\hat{\beta}_{ave}) = \frac{\iota' \Omega \iota}{S} \geq w' \Omega w = \mathrm{var}(\beta^*_w) \tag{17}$$

That is, the optimal regression in the class (9) is one of $y_i^* = \sum_{s=1}^{S} w_s y_{is}$ on $\bar{x}_i$, rather than that of $\bar{y}_i$ on $\bar{x}_i$.

4. An application

In order to examine how far the weights dictated by the minimum variance estimator in the class (9) differ from a rule of equal weighting, we look at income 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 (1993).

We have averaged the incomes of parents over the five-year period 1967-71, and we have data on the incomes of children for the three consecutive years 1987, 1988, 1989. The income concept taken here is total family income (measured in 1967 dollars–the year prior to the which the survey was started). Parents and children are at least 25 years of age when their incomes are observed, and we have selected one child per parent family. Finally, as income is bound to vary over the life cycle in a non-random way, we have run prior regressions of the logarithm of 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.

Table 1 about here

The range of estimates of the intergenerational correlation using $\hat{\delta}_s$ for the three consecutive years 1987, 1988 and 1989 is perhaps larger than one would expect: β is estimated at 0.379 using the 1987 wave, at 0.416 using the 1988 data, and at 0.435 the following year. Estimates of the signal to total variance ratio $\sigma_{pp}/(\sigma_{pp} + \sigma_{\varepsilon\varepsilon})$ for the ordinary least squares estimator of β are approximately in the range of 0.65 to 0.80, (see for instance Bowles, 1972 and Zimmerman, 1992).

If say $\sigma_{pp} = 3/4$ and $\sigma_{\varepsilon\varepsilon} = 1/4$ implying a signal to total variance ratio of 3/4, averaging the incomes of parents over a 5-year period, as we have done here, would still depress the estimate of β by 6.7%. That is, multiplying our estimates $\hat{\delta}_s$ by this hypothetical correction factor would entail an even larger range of values, 0.404 to 0.464, for the intergenerational correlation.

Averaging the incomes of children (or equivalently, as we have seen, averaging the various $\hat{\delta}_s$'s) substantially reduces the range of estimates of β. There are four estimates reported in table 1: three estimates taking two years of income for the child at a time, and an average of all three years. The Behrman and Taubman variant, $\hat{\beta}_{ave}$, narrows the range of estimates to [0.397; 0.425]. Applying our hypothetical correction factor of 1.067, we would be led to conclude that β is perhaps more in the [0.424; 0.453] range, rather than the [0.404; 0.464] interval mentioned earlier.

The estimator β^*_w tells us however that the minimum variance regression is one of $y^*_i = \sum_{s=1}^{S} w_s y_{is}$ on $\bar{x}_i$. Though the weights are not solely dependent on the variances of the estimators (see equation 16), it can be noted that $\hat{\delta}_{87}$ exhibits the largest standard error (and is the smallest of the three estimates in this application). Accordingly, it receives the least weight in the three cases where it enters the computation of β^*_w. $\hat{\delta}_{87}$ is a given a 0.281 weight in the 87&88 regression, a 0.333 weight in the 87&89 regression, and a 0.159 weight in the regression that utilizes all three years. The differences in weights can be seen to be much smaller when pooling $\hat{\delta}_{87}$ and $\hat{\delta}_{88}$ whose standard errors are broadly similar. The effect of adopting β^*_w here is to revise upward the $\hat{\beta}_{ave}$ estimate when the 1987 data are used. Though this conclusion could quite conceivably be reversed if the least noisy estimate $\hat{\delta}_s$ were also the smallest one, the minimum variance estimator β^*_w provides a less ad-hoc way of summarizing data over the practice of selecting the smallest, or largest, estimate of β depending on one's belief that intergenerational mobility is high, or low, in the United States and elsewhere.

5. Conclusions

Averaging methods have frequently been used in the analysis of intergenerational income continuities as a means of limiting the effect of biases resulting from measurement error. The Solon/Zimmerman variant of the method of averaging regresses a single year measure of the child's income on a T-period average of the parent family's resources, while the Behrman and Taubman estimator regresses an S-period average of the child's income on the same T-period average of the parent family's resources. The estimator resulting from the Behrman and Taubman methodology was shown to be an arithmetic mean of the S distinct slope estimates resulting from the Solon/Zimmerman variant of the method of averaging.

Because sampling variation produces year to year changes in the variance of children's incomes, it was shown that the Behrman and Taubman estimator is not e¢cient in the class of estimators which can be expressed as a weighted sum of the S slope estimates resulting from the Solon/Zimmerman methodology. The minimum variance estimator in the above class was derived, and when applied to a US sample of parents and children from the PSID, attributed a lower weight to a Solon/Zimmerman estimate with a markedly higher standard error.

6. References

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.

Fraser D. (1976): Probability and Statistics: Theory and Applications, North Scituate MA, Duxbury Press.

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

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

Solon G. (1992): "Intergenerational Income Mobility in the United States", American Economic Review 82, 393-408.

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

Table 1 : estimation results

year *s*	$\bar{b}_{87}$	$\bar{b}_{88}$	$\bar{b}_{89}$	$\hat{b}_{ave}$	b^*_w	w_{87}	w_{88}	w_{89}
1987 & 1988	0.379 (0.055)	0.416 (0.050)		0.397 (0.050)	0.405 (0.049)	0.281	0.719	
1988 & 1989		0.416 (0.050)	0.435 (0.051)	0.425 (0.048)	0.424 (0.048)		0.558	0.442
1987 & 1989	0.379 (0.055)		0.435 (0.051)	0.407 (0.050)	0.416 (0.050)	0.333		0.667
1987&88&89	0.379 (0.055)	0.416 (0.050)	0.435 (0.051)	0.410 (0.048)	0.417 (0.048)	0.159	0.472	0.369

Notes

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

2 The parents' income is averaged over the 5-year period 1967-71. The child's income is as defined in the first column of the table.

3 Standard errors appear inside parentheses, n=592.