Persistence bias and the wage-schooling model

Corrado Andini

Universidade da Madeira, Campus da Penteada, 9000-390 Funchal, Portugal
Centro de Estudos de Economia Aplicada do Atlântico (CEEAplA), 9501-801 Ponta Delgada, Portugal
Institute for the Study of Labor (IZA), D-53072 Bonn, Germany

ABSTRACT

A well-established empirical literature suggests that individual wages are persistent. Yet, the standard human-capital wage model does not typically account for this stylized fact. This paper investigates the consequences of disregarding earnings persistence when estimating a standard wage-schooling model. In particular, the problems related to the estimation of the schooling coefficient are discussed.

JEL Classification: C23, I21, J31
Keywords: schooling, wages, dynamic panel-data models

Corresponding author:

Prof. Corrado Andini
Universidade da Madeira
Campus da Penteada
9000-390 Funchal
Portugal
E-mail: andini@uma.pt
Tel.: +351 291 70 50 53
Fax: +351 291 70 50 49

This version: September 2013
An earlier version of this manuscript has been published as an IZA discussion paper in January 2013.

1. Introduction

Since the publication of a seminal article by Griliches (1977), it is known that the ordinary least squares estimator of the schooling coefficient in a simple static wage-schooling model is biased. In particular, Griliches pointed out that the least squares estimation of the schooling coefficient is subject to two types of bias, which are sometimes referred as the Griliches's biases. The first, known as the ability bias, is an upward bias due to the correlation between individual unobserved ability and schooling. The second, known as the attenuation bias, is a downward bias due to measurement errors in the schooling variable.

Attempts to cure (reduce) the Griliches's biases have been based on three main empirical approaches: extensions of the control set (to proxy unobserved error components and thus reduce the 'importance' of the error term), instrumental-variable estimation (to control for endogeneity), and the use of better data (such as longitudinal data, to control for individual unobserved heterogeneity). Of course, combinations of these approaches have also been adopted.

One striking feature of the existing literature is that the body of evidence is vast. This partly explains why it is difficult to make a definitive statement about the magnitude of the schooling coefficient, with and without correcting for the Griliches's biases. However, one of the things that we know is that, as argued by Card (2001), instrumental-variable estimates of the schooling coefficient are typically found to be bigger than least squares estimates, and more imprecise.

This paper investigates the consequences of a new (some may say old) type of bias affecting the least squares estimation of the schooling coefficient in a simple wage-schooling model. While there are hundreds of studies dealing with the Griliches's biases, to the best of our knowledge, no research has been so far conducted to highlight another important source of distortion, the bias arising from the least squares estimation of the schooling coefficient in a static wage-schooling model which disregards earnings persistence. Let us refer to it as the 'least squares persistence bias'.

The first key issue in this paper is thus whether the persistence of earnings is important or not in a model for individual wages. Obviously, disregarding earnings persistence in wage-schooling models would not cause any problem if earnings persistence were not important in individual wage models. At opposite, if earnings persistence were important, then disregarding such persistence would be problematic.

As a matter of fact, the empirical evidence on the persistent nature of earnings, both at micro and macro level, is already large. Indeed, it has already been reviewed, among others, by both Taylor (1999), who has focused on the macroeconomic evidence, and Guvenen (2009), who has instead discussed most of the existing microeconomic studies.

Focusing on the microeconomic evidence, which is particularly relevant for individual wage-schooling models, it is worth noting that the discussion about the persistence of individual wages is not new. In contrast, it dates several decades back. For instance, some of the first articles taking the dynamic aspects of individual earnings models into account have been authored in the 1970s and the 1980s by Lillard and Willis (1978), MaCurdy (1982) and Abowd and Card (1989), among others. More recently, individual-level dynamic wage models taking the persistent nature of earnings into account have been proposed and estimated by Guiso et al. (2005), Cardoso and Portela (2009), and Hospido (2012), to cite a few.

However, despite the existing empirical evidence on the persistence of individual wages, the incorporation of the persistent nature of individual earnings in human-capital or Mincerian-type models has been slow. One explanation for this fact is that it is uneasy to account for earnings persistence, endogeneity, individual unobserved heterogeneity and selection, all at the same time, even if the wage-schooling model is assumed to be linear. Nevertheless, the existing literature includes a couple of exceptions.

In particular, the importance of accounting for earnings persistence in wage-schooling models has been repeatedly stressed by Andini (2007; 2009; 2010; 2013a; 2013b). For instance, Andini (2009; 2013a) has proposed a simple theoretical model to explain why past wages should play the role of additional explanatory variable in human-capital regressions. The intuition is that, in a world where bargaining matters, the past wage of an individual can affect his/her outside option and thus the bargained current wage. Analogously, Andini (2010; 2013b) has proposed an adjustment model between observed earnings and potential earnings (the latter being defined as the monetary value of the individual human-capital productivity) where the adjustment speed is allowed to be not perfect. In addition, Andini (2013a; 2013b) has built a bridge between the literature on earnings dynamics (Guvenen, 2009) and the Mincerian literature, showing how to obtain a consistent GMM-SYS estimate of the schooling coefficient in a Mincerian wage equation when earnings persistence, endogeneity and individual unobserved heterogeneity are taken into account. Similarly, Semykina and Wooldridge (2013) have estimated a wage-schooling model accounting for earnings persistence and sample selection. Finally, Kripfganz and Schwarz (2013) have estimated a dynamic wage-schooling model using an econometric approach alternative to the GMM-SYS estimation approach suggested by Andini (2013a; 2013b).

Based on the above mentioned empirical micro evidence, this paper starts from the assumption that controlling for earnings persistence is potentially important in individual wage-schooling models. And, starting from this assumption, it elaborates on the consequences of disregarding the dynamic nature of the wage-schooling link in the least squares estimation of the schooling coefficient. In addition, despite the initial reference to the least squares estimator, the paper goes beyond that specific case by discussing the problems of other estimators. Indeed, it will be argued that the persistence bias is a general problem associated with the estimation of the schooling coefficient in a static wage-schooling model, regardless of the estimator used. For instance, it will be argued that the standard static instrumental-variable estimation is unable to solve the persistence-bias problem, i.e. there exists an 'instrumental-variable persistence bias'.

Specifically, this paper provides the following five novel findings. First, it provides an expression for the bias of the least squares estimator of the schooling coefficient in a simple wage-schooling model where earnings persistence is not accounted for. It is argued that the least squares estimator of the schooling coefficient is biased upward, and the bias is increasing with potential labor-market experience (age) and the degree of earnings persistence. Second, data from the National Longitudinal Survey of Youth (NLSY) are used to show that the magnitude of the least squares persistence bias is non-negligible. Third, the least squares persistence bias cannot be cured by increasing the control set. Fourth, an expression for the persistence bias of the standard instrumental-variable estimator of the schooling coefficient in a static wage-schooling model is provided. Finally, it is shown that disregarding earnings persistence is still problematic for the estimation of the schooling coefficient even if individual unobserved heterogeneity and endogeneity are taken into account. The case of the Hausman-Taylor estimator is considered.

In short, the standard cures for the Griliches's biases (based on extensions of the control set, treatments of endogeneity and models with individual unobserved heterogeneity) are unable to solve the persistence-bias problem related to the estimation of static wage-schooling models. Therefore, an enormous number of schooling coefficient estimates, based on static models, is potentially subject to the persistence-bias critique.

Overall, the findings support the dynamic approach to the estimation of wage-schooling models recently suggested by Andini (2013a; 2013b).

The rest of the paper is organized as follows. Section 2 provides an expression for the persistence bias of the least squares estimator for the schooling coefficient. Section 3 investigates the magnitude of that bias using US data on young male workers. Section 4 analyzes whether the bias can be somehow reduced by extending the control set. Section 5 provides an expression of the persistence bias of the standard instrumental-variable estimator for the schooling coefficient. Section 6 highlights that disregarding earnings persistence is still problematic even if individual unobserved heterogeneity and endogeneity are accounted for. In particular, the case of the Hausman-Taylor estimator is discussed. Section 7 concludes.

2. Persistence bias in static least squares models

This section provides an expression for the persistence bias of the least squares estimator of the schooling coefficient, under a set of simplifying hypotheses.

Let us consider a simple wage-schooling model. In particular, let us assume that the 'true' model is as follows:

$$(1) \quad w_{i,s+z+1} = \alpha + \beta s_i + \rho w_{i,s+z} + u_{i,s+z+1} \quad \text{for } \forall i, s+z \quad \text{with } s \geq 1 \;\; z \geq 0$$

where w is logarithm of gross hourly wage, s is schooling years, z is years of potential labor-market experience, and u is an error term[1]. Hence the 'true' model is dynamic in the sense that past wages help to predict current wages.

In addition, let us assume that:

[1] Following the standard Mincerian model, it is assumed that an individual starts working after leaving school. The first observed wage is observed in year s.

(A1) $COV(s_i, u_{i,s+z+1}) = 0 \qquad \forall i, s+z$

(A2) $COV(w_{i,s+z}, u_{i,s+z+1}) = 0 \qquad \forall i, s+z$

(A3) $COV(u_{i,s+z}, u_{i,s+z+1}) = 0 \qquad \forall i, s+z$

(A4) $COV(u_{i,s+z}, u_{j,s+z}) = 0 \qquad \forall i \neq j, s+z$

(A5) $E(u_{i,s+z+1}) = 0 \qquad \forall i, s+z$

(A6) $VAR(u_{i,s+z+1}) = \theta^2 \qquad \forall i, s+z$

(A7) $VAR(s_i) = \sigma^2 \qquad \forall i$

(A8) $COV(s_i, \rho w_{i,s-1} + u_{i,s}) = 0 \qquad \forall i, s$

Assumption (A1) basically means that we exclude the Griliches's biases in order to focus on the persistence bias. Assumption (A2) is an additional condition required for the least squares estimator of model (1) to be consistent (it excludes the so-called Nickell's bias). Of course, both these assumptions are unlikely to hold. However, we will discuss the implications of removing them later on in this paper. First, we will use these simplifying assumptions to make the first point of this paper, which is about the inconsistency of the least squares estimator for the schooling coefficient when the wage-schooling model does not take into account earnings persistence.

Assumptions from (A3) to (A7) are quite standard. Assumption (A8), instead, is not standard. It can be seen as an 'initial condition'. One may think at $w_{i,s-1}$ as a reservation wage[2] that every individual has in mind before leaving school, at time $s-1$.

[2] The idea of a reservation wage is compatible with the presence of self-selection into the labor market. However, in this paper, we do not deal with this important issue. We just consider the estimation of a wage equation where earnings persistence, individual unobserved heterogeneity and endogeneity matter (see also footnote 5).

Yet, this wage is not observed. Hence, at time s, the error term in model (1) will be given by $(\rho w_{i,s-1} + u_{i,s})$. It may well be the case that this reservation wage is correlated with s_i as higher educated people are likely to have higher reservation wages. However, assumption (A8) excludes this possibility. The reason is simple and related to assumption (A1): at this stage, in order to focus on the least squares persistence bias, we exclude all sources of bias due to correlation between schooling and the error term in model (1). Again, we will discuss the implications of removing these simplifying assumptions later on.

Under the above hypotheses, a proof of the inconsistency of the least squares estimator applied to a simple static wage-schooling model is straightforward. In short, if the 'true' model is (1) but earnings persistence is disregarded and the following static 'false' model is estimated:

(2) $$w_{i,s+z+1} = \alpha + \beta s_i + e_{i,s+z+1} \qquad \text{where } e_{i,s+z+1} = \rho w_{i,s+z} + u_{i,s+z+1}$$

then, it is easy to show that:

(3) $$\text{p}\lim \beta_{OLS} = \beta + \rho \frac{COV(s_i, w_{i,s+z})}{VAR(s_i)}$$

Knowing that $VAR(s_i) = \sigma^2$, it is possible to focus on $COV(s_i, w_{i,s+z})$. In particular, it can be shown that:

(4) $$\begin{aligned} &COV(s_i, w_{i,s+z}) = COV(s_i, \alpha + \beta s_i + \rho w_{i,s+z-1} + u_{i,s+z}) = \\ &= \beta\sigma^2 + \rho COV(s_i, w_{i,s+z-1}) = \beta\sigma^2 + \rho COV(s_i, \alpha + \beta s_i + \rho w_{i,s+z-2} + u_{i,s+z-1}) = \\ &= \beta\sigma^2 + \rho\left[\, \beta\sigma^2 + \rho COV(s_i, w_{i,s+z-2}) \,\right] = \beta\sigma^2 + \rho\beta\sigma^2 + \rho^2 COV(s_i, w_{i,s+z-2}) = \\ &= \beta\sigma^2(1 + \rho + \rho^2 + ... + \rho^{z-1}) + \rho^z COV(s_i, w_{i,s}) \end{aligned}$$

Since $COV(s_i, w_{i,s}) = COV(s_i, \alpha + \beta s_i + \rho w_{i,s-1} + u_{i,s}) = \beta\sigma^2 + COV(s_i, \rho w_{i,s-1} + u_{i,s})$ and $COV(s_i, \rho w_{i,s-1} + u_{i,s}) = 0$ by assumption, then we get:

(5) $$\begin{aligned} COV(s_i, w_{i,s+z}) &= \beta\sigma^2(1+\rho+\rho^2+...+\rho^{z-1}) + \rho^z\beta\sigma^2 = \\ &= \beta\sigma^2(1+\rho+\rho^2+...+\rho^z) \end{aligned}$$

Hence, using (3), it follows that:

(6) $$p\lim\beta_{OLS} = \beta + \rho\beta\sum\rho^z$$

where $\rho\beta\sum\rho^z$ is the persistence bias. The conclusion is that the least squares estimator of the schooling coefficient in model (2) is biased upward if β and ρ are positive, with the bias increasing in both ρ and z.

As a matter of example, Figure 1 illustrates how the bias increases with z using $\rho = 0.600$ and $\beta = 0.030$ as simulation parameters.

3. Is the persistence bias worrisome in static least squares models?

It is interesting to discuss the magnitude of the persistence bias when estimating a simple static wage-schooling model with real data. Particularly, we find of interest to explore a well-known publically available dataset of US young workers, in which the persistence bias should be lower than in a standard dataset including older workers since the average potential experience (z) is lower (as the average age is lower).

Specifically, in this paper, the data are taken from the National Longitudinal Survey of Youth (NLSY). The dataset contains observations on 545 males for the period of 1980-1987. To our knowledge, this dataset has been already used by Vella and Verbeek (1998), Wooldridge (2005) and Andini (2007; 2013a), among others.

The summary statistics of the variables and their meaning are presented in Appendix. The dataset has four main advantages: it is a balanced panel (which avoids a number of econometric issues with unbalanced panels), it is publically available (making replication easier), it has been already used in the literature (making comparison with earlier studies possible) and it has already been cleaned up, such that the schooling variable is actually time-invariant.

The estimation results, obtained using the least squares estimator, are presented in Table 1. Column 1 presents the least squares estimates from model (1), the ‘true’ dynamic one. The coefficient of schooling β is estimated at 0.034, with the degree of earnings persistence ρ estimated at 0.599. Column 2 provides the estimate of the schooling coefficient from the 'false' static model (2), which does not control for earnings persistence. As expected, the estimate of the schooling coefficient is well above the 'true' value of the coefficient. Indeed, the coefficient is estimated at 0.076. The difference between 0.076 and 0.034 can be seen as a proxy of the persistence bias, under Section 2's assumptions. Since the average potential experience (z) in the sample is 6.5 years, a 0.042 bias is perfectly in line with our theoretical prediction in Section 2 (see Figure 1), and its magnitude is non-negligible.

4. Does extending the control set cure the persistence bias in static least squares models?

Columns 3 to 7 gradually extend the static model (2) to investigate whether the persistence bias can be somehow cured (reduced) by increasing the control set, i.e. by improving the explanatory power of the static model (2) and searching for 'substitutes' of the past wage.

For instance, column 3 proposes the classical Mincerian specification which controls for potential experience and its square. However, the coefficient of schooling does not decrease, thus indicating that potential experience (age) is not a substitute for past wage. In contrast, the schooling coefficient increases to 0.102.

Columns from 4 to 7 add a number of individual specific characteristics, both time-varying and constant, which increase the explained variability of wages, though not as much as just controlling for past wage, as the evolution of the R-squared coefficient suggests. In particular, column 4 takes into account union membership, marital status, public-sector employment, race (whether the individual is Black or Hispanic) as well as presence of health disabilities. Column 5 adds information on the individual residence (whether the individual lives in the South, Northern Central or North East). In addition, it controls for whether the individual lives in a rural area or not. Columns 6 and 7 add detailed information on industry and occupation, respectively.

The key point in this section is that no static specification is able to provide a coefficient of schooling close to the 'true' one, estimated using model (1).

Finally, column 8 adds year fixed effects to model (2). They are found to be not jointly significant (p-value 0.232). In addition, the R-squared coefficient does not significantly improve. Hence, likewise the experience variables, year effects cannot be seen as substitutes for past wage. At best, year effects can be seen as substitutes for experience variables themselves. However, to keep a Mincerian-type specification, in the rest of this paper, we will continue keeping experience variables in the control set, thus excluding year effects. Hence, our full control set will be the one used in column 7.

5. Persistence bias in static instrumental-variable models

So far, we have focused our attention on the least squares estimator. Yet, as it is well known, the estimate of the schooling coefficient in model (1) based on the least squares estimator cannot be taken as a good proxy of the ‘true’ value of the schooling parameter due to the correlation between errors and schooling (the Griliches’s biases) and/or between errors and lagged wage (the Nickell’s bias). Such correlation causes the least squares estimator of model (1) to be inconsistent.

To fix the ideas, let us assume that the error term $u_{i,s+z+1}$ in model (1) would be better seen as the sum between individual-specific unobserved effects c_i, representing individual abilities or measurement errors in the schooling variable, and a ‘well-behaved’ disturbance $v_{i,s+z+1}$. That is, let us assume that $u_{i,s+z+1} = c_i + v_{i,s+z+1}$ with[3]:

(A9) $COV(s_i, c_i) \neq 0$ $\qquad \forall i$

(A10) $COV(s_i, v_{i,s+z+1}) = 0$ $\qquad \forall i, s+z$

(A11) $COV(w_{i,s+z}, v_{i,s+z+1}) = 0$ $\qquad \forall i, s+z$

By introducing individual-specific unobserved effects, we introduce two sources of bias for the least squares estimator applied to model (1). The first one is assumption (A9) which implies that assumptions (A1) and (A8) do not hold true any more. The second one is that $w_{i,s+z}$ turns out to be correlated with c_i, making assumption (A2) invalid.

[3] For simplicity, we avoid reporting a couple of additional standard assumptions about $v_{i,s+z+1}$.

The literature has typically dealt with the violation of assumption (A1) using instrumental variables. However, while a big research effort has been oriented towards the search of the best instrumental variable, the presence of the past wage in model (1) has been generally neglected. Indeed, the standard practice has been to estimate the 'false' static model, i.e. model (2), under the implicit assumption that $e_{i,s+z+1} = \rho w_{i,s+z} + u_{i,s+z+1}$ and $u_{i,s+z+1} = c_i + v_{i,s+z+1}$. The key point of this section is precisely that the standard practice has been, in fact, incorrect because disregarding the past wage biases the instrumental-variable estimation of the schooling coefficient in model (2).

A simple proof of why a static instrumental-variable approach can be misleading is as follows. Let us suppose that a researcher worries about a possible correlation between $u_{i,s+z+1}$ and s_i, but the role played by the past wage in model (1) is disregarded. In short, the researcher assumes that $\rho = 0$ while this hypothesis does not hold true. The standard static instrumental-variable practice is to find a time-invariant instrument g_i such that $COV(g_i, s_i) \neq 0$ (for instance, the schooling years of the father of the individual i). In this case, it is easy to show that:

$$(7) \qquad p\lim \beta_{IV} = \beta + \frac{COV(g_i, u_{i,s+z+1})}{COV(g_i, s_i)} + \frac{COV(g_i, \rho w_{i,s+z})}{COV(g_i, s_i)}$$

The conclusion is that, even if the researcher is able to find an instrument satisfying $COV(g_i, u_{i,s+z+1}) = 0$, i.e. the standard instrumental-variable assumption, the instrumental-variable estimator will still be inconsistent[4] as $COV(g_i, s_i) \neq 0$ implies $COV(g_i, \rho w_{i,s+z}) \neq 0$. This is trivial because $w_{i,s+z}$ is correlated with s_i.

[4] Another source of bias for the instrumental-variable estimator in static models is the presence of heterogeneous returns to schooling, i.e. the case in which the schooling coefficient is not the same across individuals. There is a rapidly-growing body of literature on this topic with recent important contributions by Carneiro, Heckman and Vytlacil, among others. In this paper, we have not explored the intersection between heterogeneous returns and earnings persistence. However, the latter is an interesting topic for future research.

This instrumental-variable inconsistency result, based on a persistence-bias critique, appears to be of fundamental importance due to its implications for the standard static approach in the Mincerian or human-capital literature. In addition, it is also important for the experimental literature since, as stressed by Carneiro et al. (2006, p. 2), the instrumental-variable method "is the most commonly used method of estimating β. Valid social experiments or valid natural experiments can be interpreted as generating instrumental variables". Yet, the autoregressive nature of wages is typically not taken into account in the experimental literature.

6. Persistence bias in static Hausman-Taylor (panel data) models

This section argues that disregarding earnings persistence is still problematic for the estimation of the schooling coefficient even if individual unobserved heterogeneity and endogeneity are taken into account. We will show that the persistence bias is a problem related to the estimation of a static wage-schooling model, regardless of whether this estimation is performed using an estimator which exploits the longitudinal structure of the dataset and takes both unobserved heterogeneity and endogeneity into account.

To make the point of this section, borrowing from Andini (2013a; 2013b), we will first present a method to obtain consistent estimates of both the schooling coefficient and the degree of earnings persistence when individual unobserved heterogeneity, endogeneity and earnings persistence are taken into account. The method is based on the GMM-SYS estimator developed by Blundell and Bond (1998). Afterwards, we will focus on the distortion of the least squares estimator, which takes into account earnings persistence but disregards both unobserved heterogeneity and endogeneity. Finally, we will discuss the main point of this section by considering the Hausman-Taylor estimator, which takes into account unobserved heterogeneity and endogeneity but disregards earnings persistence.

6.1 How to obtain consistent estimates: the GMM-SYS estimator

Under the new reasonable assumptions made in Section 5, Andini (2013a; 2013b) has shown that consistent[5] estimates for ρ and β are obtained using the GMM-SYS

[5] One limitation of the approach proposed by Andini (2013a; 2013b) is that selection is not considered. A dynamic wage-schooling model where selection matters has been estimated by Semykina and Wooldridge (2013). Yet, in their approach, a non-zero correlation between the time-constant variables and time-invariant individual

estimator proposed by Blundell and Bond (1998), i.e. by using the following equations as a system:

(8) $\Delta w_{i,s+z+1} = \rho \Delta w_{i,s+z} + \Delta v_{i,s+z+1}$

(9) $w_{i,s+z+1} = \alpha + \beta s_i + \rho w_{i,s+z} + c_i + v_{i,s+z+1}$

and using $w_{i,s+z-1}$ and $\Delta w_{i,s+z-1}$ as instruments for (8) and (9), respectively.

Of course, the use of $\Delta w_{i,s+z-1}$ and further lags as instruments is the key assumption to identify the schooling coefficient and it has the advantage to be easily testable. In particular, the additional orthogonality conditions imposed by the level equation (9) must pass the Difference-in-Hansen test.

A further requirement is that the level-equation instruments should not be weak. This may happen in presence of non-stationary variables. The latter is also an easily testable assumption. A test can be based on the estimation of an AR1 process (with constant term) for the variable in levels, again using the GMM-SYS estimator. A preliminary test can be based on the least squares estimator, which typically overestimates the autoregressive coefficient (see Blundell and Bond, 2000). For instance, in our sample, using the least squares estimator, the autoregressive coefficient of the AR1 log-wage process (with constant term) is estimated at 0.626 with robust standard error of 0.025 and p-value equal to 0.000. Hence, it is likely that the true autoregressive coefficient of the log-wage process is well below the critical value of 1.000.

Of course, if one or more variables are found to be non-stationary, they should be excluded from the set of level-equation instruments.

Using the full control set, the GMM-SYS estimator provides an estimate of the degree of earnings persistence ρ equal to 0.174 and an estimate of the schooling coefficient β equal to 0.102, both significant at 1% level.

6.2 Bias in dynamic least squares models

---

unobserved heterogeneity implies that the effect of time-constant observed variables, such as schooling, cannot be distinguished from that of the unobserved heterogeneity (Semykina and Wooldridge, 2013, p. 50).

Taking the above estimates as the ‘true’ values of the corresponding parameters, it is interesting to discuss the biases implied by alternative estimators or models, with special attention to the coefficient of schooling.

The first thing to note is that Andini (2013b) has already investigated the consequences for the least squares estimator of removing assumptions (A1), (A2) and (A8). In particular, using Belgian data, the author has pointed to an upward-biased estimate of the degree of earnings persistence and to a downward-biased estimate of the schooling coefficient.

Estimation with NLSY data in Table 2 confirms the above view. Column 1 reports the least squares estimates of model (1) with no controls. Column 2 adds all the controls considered in column 7 of Table 1, i.e. the full control set. The finding is that there is no big difference in the estimates of both β and ρ between column 1 and column 2. However, once individual unobserved heterogeneity and endogeneity are taken into account using the GMM-SYS estimator, the finding is different. Indeed, column 3 shows that the least squares estimator, used in column 2 (and column 1), seems to overestimate the degree of earnings persistence and to underestimate the schooling coefficient. So, the problem with the least squares approach to model (1) is that it does not take into account individual unobserved heterogeneity and endogeneity.

6.3 Persistence bias in static panel data models

Yet, the key point in this section is not about the failure of dynamic least squares models. The key point here is to highlight how misleading can be the static-model estimation of the schooling coefficient, even when the control set is large and when both individual unobserved heterogeneity and endogeneity are taken into account. To this end, Table 3 presents some additional evidence comparing the ‘true’ estimate of the schooling coefficient based on the GMM-SYS estimator, again reported in column 3, with an estimate based on an important instrumental-variable estimator for static panel data models.

In particular, we consider an estimator which is typically used when time-invariant variables, such as schooling, are included in the explanatory set: the Hausman-Taylor estimator. As a benchmark, we also report an estimate of the schooling coefficient based on a different estimator for static panel data models: the Random Effects estimator.

The Random Effects estimator, used in column 1 of Table 3, exploits the longitudinal nature of the dataset and controls for individual unobserved effects under the assumption that they are uncorrelated with schooling and other explanatory variables. The Hausman-Taylor estimator, used in column 2, additionally takes into account that some explanatory variables, including schooling, can be endogenous. Hence, the Hausman-Taylor estimator takes both individual unobserved heterogeneity and endogeneity into account, but it disregards earnings persistence.

In all the columns of Table 3, the control set used is the full one. In particular, in the Hausman-Taylor estimation, HLTH is taken as time-varying exogenous, BLACK and HISP are taken as time-invariant exogenous, SCHOOL is taken as time-invariant endogenous, and all the other variables in the full control set are taken as time-varying endogenous. The identification is based on the standard Hausman-Taylor approach. For instance, the mean value of HLTH is used as instrument for SCHOOL.

Focusing on the Hausman-Taylor estimation, the conclusion seems to be that again, likewise the classical instrumental-variable case, disregarding earnings persistence can be problematic. Indeed, the coefficient of schooling based on the Hausman-Taylor estimator (0.220) more than doubles the ‘true’ one (0.102). This is the key result of the comparison between column 2 and column 3 in Table 3.

The good news for static-model users is that the GMM-SYS estimate of the schooling coefficient seems to be in line with the Random Effects estimate (0.090). This can be observed by comparing column 1 and column 3 in Table 3. More interestingly, the static least squares Mincerian model in column 3 of Table 1 seems to provide a very good proxy for the ‘true’ coefficient (0.102), suggesting that, once a quadratic function of experience is accounted for, the least squares estimator may benefit from the possibility that persistence, ability, attenuation and omitted-variable biases compensate each other. Although we are sceptical about the possibility of such a compensation to be systematic, we believe that this finding is something worth mentioning.

6. Conclusions

There are at least three intuitive reasons why wage-schooling models should by handled as dynamic models: i) individual human-capital productivity and wages may not adjust instantaneously due to frictions in the labour market (Andini, 2010; 2013b); ii) past wages may affect the outside option of an individual in a simple bargaining model over wages and productivity (Andini, 2009; 2013a); iii) the residuals of the wage equation, representing wage or productivity shocks, may be show some degree of persistence (Guvenen, 2009, among many others, models them as autoregressive of order one). Of course, combinations of these explanations enrich the set of possibilities.

Despite the above theoretical arguments and an already large body of evidence supporting the dynamic behaviour of individual wages, the existing human-capital literature has not paid sufficient attention to the dynamic nature of the link between schooling and wages. Indeed, while examples of the estimation of static versions of the wage-schooling model are abundant, examples of estimated dynamic wage-schooling models can be counted on the fingers of one hand.

This pattern of the human-capital literature, however, should not be surprising. The initial theoretical wage-schooling models put forward by the fathers of modern education economics (Becker, Ben-Porath and Mincer, to cite a few) were particularly clever and their predictions have inspired a large body of static model evidence. In addition, longitudinal datasets including information on individual characteristics have not been easily accessible for several decades, making dynamic micro-level empirical analyses not executable. Fortunately, at least with respect to the latter aspect, today's reality is different. Longitudinal datasets are freely available and the issue raised in this paper can now receive the appropriate consideration from the research community. Whether this will happen or not is still an open question.

Starting from the above motivation, this paper has investigated the consequences of disregarding earnings persistence when estimating a standard wage-schooling model. We have argued that the estimation of the schooling coefficient in a static wage-schooling model is, in general, biased.

Five main results have been presented in this paper. First, the least squares estimator of the schooling coefficient has been shown to be biased upward, with a bias increasing in potential labor-market experience (age) and the degree of earnings persistence. Second, the least squares persistence bias has been found to be non-negligible in NLSY data. Third, the least squares persistence bias has be found to be non-curable by increasing the control set. Fourth, the standard static instrumental-variable approach has been shown to be inconsistent. Finally, disregarding earnings persistence has been argued to be still problematic even when the estimator used accounts for individual unobserved heterogeneity and endogeneity. The case of the Hausman-Taylor estimator has been discussed.

Overall, the findings support the dynamic approach to the estimation of wage-schooling models recently proposed by Andini (2013a; 2013b).

References

Abowd, J., Card, D. (1989) On the covariance structure of earnings and hours changes. Econometrica, 57(2): 411-445.

Andini, C. (2007) Returns to education and wage equations: A dynamic approach. Applied Economics Letters, 14(8): 577-579.

Andini, C. (2009) Wage bargaining and the (dynamic) Mincer equation. Economics Bulletin, 29(3): 1846-1853.

Andini, C. (2010) A dynamic Mincer equation with an application to Portuguese data. Applied Economics, 42(16): 2091-2098.

Andini, C. (2013a) How well does a dynamic Mincer equation fit NLSY data? Evidence based on a simple wage-bargaining model. Empirical Economics, 44(3): 1519-1543.

Andini, C. (2013b) Earnings persistence and schooling returns. Economics Letters, 118(3): 482-484.

Blundell, R.W., Bond, S.R. (1998) Initial conditions and moment restrictions in dynamic panel data models. Journal of Econometrics, 87(1): 115-143.

Blundell, R.W., Bond, S.R. (2000) GMM estimation with persistent panel data: an application to production functions. Econometric Reviews, 19(3): 321-340.

Card, D. (2001) Estimating the return to schooling: progress on some persistent econometric problems. Econometrica, 69(5): 1127-1160.

Cardoso, A., Portela, M. (2009) Micro foundations for wage flexibility: wage insurance at the firm level. Scandinavian Journal of Economics, 111(1): 29-50.

Carneiro, P., Heckman, J., Vytlacil, E. (2006) Estimating marginal and average returns to education. Unpublished manuscript available at: http://www.ucl.ac.uk/~uctppca/school_all_2006-10-30b_mms.pdf (accessed: 13 Jan. 2013)

Griliches, Z. (1977) Estimating the returns to schooling: some econometric problems. Econometrica, 45(1): 1-22.

Guiso, L., Pistaferri, L., Schivardi, F. (2005) Insurance within the firm. Journal of Political Economy, 113(5): 1054-1087.

Guvenen, F. (2009) An empirical investigation of labor income processes. Review of Economic Dynamics, 12(1): 58-79.

Hospido, L. (2012) Modelling heterogeneity and dynamics in the volatility of individual wages. Journal of Applied Econometrics, 27(3): 386-414.

Kripfganz, S., Schwarz, C. (2013) Estimation of linear dynamic panel data models with time-invariant regressors. Deutsche Bundesbank discussion paper nº 25/2013, May (last update: 30 June 2013).

Lillard, L., Willis, R. (1978) Dynamic aspects of earnings mobility. Econometrica, 46(5): 985-1012.

MaCurdy, T. (1982) The use of time-series processes to model the error structure of earnings in longitudinal data analysis. Journal of Econometrics, 18(1): 83-114.

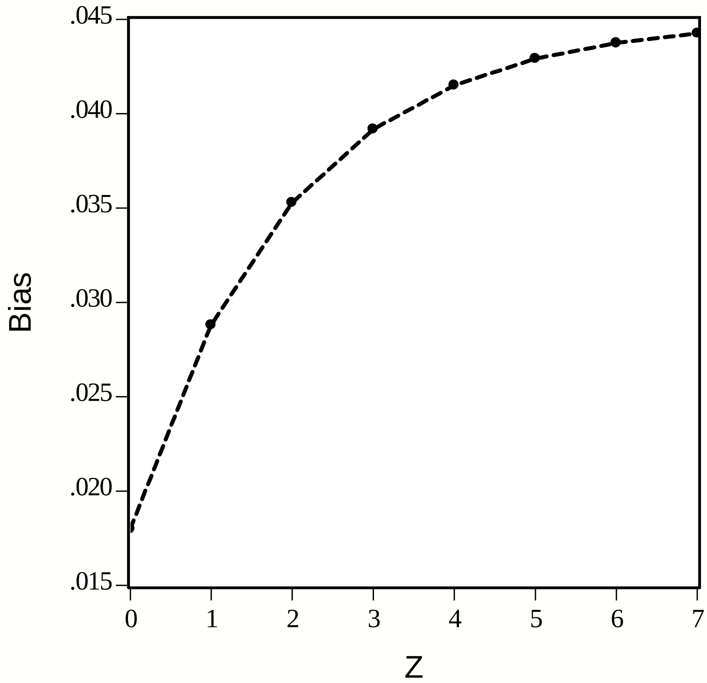
Semykina, A., Wooldridge, J.M. (2013) Estimation of dynamic panel data models with sample selection. Journal of Applied Econometrics, 28(1): 47-61.

Taylor, J.B. (1999) Staggered price and wage setting in macroeconomics. In: Taylor J.B., Woodford M. (Eds.) Handbook of Macroeconomics, New York: North Holland.

Vella, F., Verbeek, M. (1998) Whose wages do unions raise? A dynamic model of unionism and wage rate determination for young men. Journal of Applied Econometrics, 13(2): 163-183.

Wooldridge, J.M. (2005) Simple solutions to the initial conditions problem in dynamic, nonlinear panel data models with unobserved effects. Journal of Applied Econometrics, 20(1): 39-54.

Figure 1



Simulation parameters: $\rho = 0.600$ and $\beta = 0.030$

Table 1

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	OLS	OLS	OLS	OLS	OLS	OLS	OLS	OLS
	Model (1)	Model (2)	Model (2)	Model (2)	Model (2)	Model (2)	Model (2)	Model (2)
Control set			Ext 1	Ext 2	Ext 3	Ext 4	Full	Full + YE
SCHOOL	0.034***	0.076***	0.102***	0.099***	0.093***	0.090***	0.078***	0.073***
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.005)
L.WAGE	0.599***							
	(0.026)							
Observations	3,815	4,360	4,360	4,360	4,360	4,360	4,360	4,360
R-squared	0.429	0.064	0.148	0.187	0.204	0.264	0.278	0.280
Controls added to model (2) in previous column			EXPER EXPER2	UNION PUB MAR BLACK HISP HLTH	S NC NE RUR	MIN CON TRAD TRA FIN BUS PER ENT MAN PRO	OCC1 OCC2 OCC3 OCC4 OCC5 OCC6 OCC7 OCC8	YEAR80 YEAR81 YEAR82 YEAR83 YEAR84 YEAR85 YEAR86

Excluded categories: AG, OCC9 and YEAR87
Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 2

	(1)	(2)	(3)
	OLS	OLS	GMM-SYS
	Model (1)	Model (1)	Model (1)
Control set		Full	Full
SCHOOL	0.034***	0.037***	0.102***
	(0.004)	(0.004)	(0.028)
L.WAGE	0.599***	0.503***	0.174***
	(0.026)	(0.028)	(0.031)
Observations	3,815	3,815	3,815
R-squared	0.429	0.469	
IUH accounted	No	No	Yes
Endogeneity accounted	No	No	Yes
Persistence accounted	Yes	Yes	Yes
Number of individuals			545
Number of instruments			171
ABAR1 test (p-value)			0.000
ABAR2 test (p-value)			0.307
Hansen test for all instruments (p-value)			0.246
Difference-in-Hansen test for level equation (p-value)			0.178

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 3

	(1)	(2)	(3)
	RE	HT	GMM-SYS
	Model (2)	Model (2)	Model (1)
Control set	Full	Full	Full
SCHOOL	0.090***	0.220	0.102***
	(0.008)	(0.172)	(0.028)
L.WAGE			0.174***
			(0.031)
Observations	4,360	4,360	3,815
IUH accounted	Yes	Yes	Yes
Endogeneity accounted	No	Yes	Yes
Persistence accounted	No	No	Yes
Number of individuals	545	545	545
Number of instruments			171
ABAR1 test (p-value)			0.000
ABAR2 test (p-value)			0.307
Hansen test for all instruments (p-value)			0.246
Difference-in-Hansen test for level equation (p-value)			0.178

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Appendix. Sample descriptive statistics for NLSY data

The data are taken from the National Longitudinal Survey of Youth. The dataset contains observations on 545 males for the period of 1980-1987. The statistics of the variables and their meaning are as follows:

Variable	Obs	Mean	Std. Dev.	Min	Max
NR	4360	5262.059	3496.150	13	12548
YEAR	4360	1983.500	2.291	1980	1987
AG	4360	0.032	0.176	0	1
BLACK	4360	0.115	0.319	0	1
BUS	4360	0.075	0.264	0	1
CON	4360	0.075	0.263	0	1
ENT	4360	0.015	0.122	0	1
EXPER	4360	6.514	2.825	0	18
EXPER2	4360	50.424	40.781	0	324
FIN	4360	0.036	0.188	0	1
HISP	4360	0.155	0.362	0	1
HLTH	4360	0.016	0.129	0	1
MAN	4360	0.282	0.450	0	1
MAR	4360	0.438	0.496	0	1
MIN	4360	0.015	0.123	0	1
NC	4360	0.257	0.437	0	1
NE	4360	0.190	0.392	0	1
OCC1	4360	0.103	0.305	0	1
OCC2	4360	0.091	0.288	0	1
OCC3	4360	0.053	0.224	0	1
OCC4	4360	0.111	0.314	0	1
OCC5	4360	0.214	0.410	0	1
OCC6	4360	0.202	0.401	0	1
OCC7	4360	0.091	0.289	0	1
OCC8	4360	0.014	0.120	0	1
OCC9	4360	0.116	0.321	0	1
PER	4360	0.016	0.128	0	1
PRO	4360	0.076	0.265	0	1
PUB	4360	0.040	0.196	0	1
RUR	4360	0.203	0.402	0	1
S	4360	0.350	0.477	0	1
SCHOOL	4360	11.766	1.746	3	16
TRA	4360	0.065	0.247	0	1
TRAD	4360	0.268	0.443	0	1
UNION	4360	0.244	0.429	0	1
WAGE	4360	1.649	0.532	-3.579	4.051

NR	Observations number
YEAR	Year of observation
SCHOOL	Schooling years
EXPER	Potential labor-market experience
EXPER2	Experience squared
UNION	Wage set by collective bargaining
MAR	Married
BLACK	Black
HISP	Hispanic
HLTH	Has health disability
RUR	Lives in rural area
NE	Lives in North East
NC	Lives in Northern Central
S	Lives in South
WAGE	Log of gross hourly wage

Occupational dummies:

OCC1	Professional, technical and kindred
OCC2	Managers, officials and proprietors
OCC3	Sales workers
OCC4	Clerical and kindred
OCC5	Craftsmen, foremen and kindred
OCC6	Operatives and kindred
OCC7	Laborers and farmers
OCC8	Farm laborers and foreman
OCC9	Service workers

Industry dummies:

AG	Agricultural
MIN	Mining
CON	Construction
TRAD	Trade
TRA	Transportation
FIN	Finance
BUS	Business and repair services
PER	Personal services
ENT	Entertainment
MAN	Manufacturing
PRO	Professional and related services
PUB	Public Administration