Center
for
Economic Research

No. 2000-25

# ESTIMATING DYNAMIC MODELS FROM REPEATED CROSS-SECTIONS

By Marno Verbeek and Francis Vella

March 2000

ISSN 0924-7815

# Estimating Dynamic Models from Repeated Cross-Sections*

Marno Verbeek
K.U. Leuven and Tilburg University

Francis Vella
Rutgers University and New York University

February 29, 2000

JEL-codes: C21, C23, C81
Key words: repeated cross-sections; pseudo panel data; cohorts; instrumental variables; individual dynamics

**Abstract**

A major attraction of panel data is the ability to estimate dynamic models on an individual level. Moffitt (1993) and Collado (1998) have argued that such models can also be identified from repeated cross-section data. In this paper we reconsider this issue. We review the identification conditions underlying alternative estimators and present an instrumental variables type estimator that is consistent under a relatively weak set of conditions. These conditions, however, are not trivially satisfied in applied work.

*Correspondence: Marno Verbeek, K.U. Leuven, Center for Economic Studies, Naamsestraat 69, 3000 Leuven, Belgium. E-mail: Marno.Verbeek@econ.kuleuven.ac.be (corresponding author); and Frank Vella, Dept. of Economics, Rutgers University, New Jersey Hall, New Brunswick, NJ 08903-5055, USA. E-mail: vella@fas-econ.rutgers.edu.

# 1 Introduction

An important feature of panel data is its facilitation of the estimation of dynamic economic models. Recently Moffitt (1993) and Collado (1998) have argued that such models can also be estimated via repeated cross-section data. The suggested estimators are either explicitly, or implicitly, instrumental variables procedures based on the aggregation of individual data into cohorts comprising individuals with similar observed characteristic(s). The lagged dependent variable is first replaced by a predicted value from an auxiliary regression and the dynamic model is subsequently estimated via instrumental variables. This paper reviews the identification conditions for consistent estimation of a linear dynamic model from repeated cross-sections. We first show that the estimator suggested by Moffitt (1993), based on ordinary least squares (OLS) in a model where the lagged dependent variable is replaced by a predicted value, is inconsistent in the absence of strong conditions imposed on the exogenous variables. We then consider several alternative instrumental variables (IV) estimators. We show that the instruments are likely to be correlated with the equations' error term, thereby causing the resulting estimators to be inconsistent. Correcting for this is similar to allowing for "cohort specific fixed effects", as is done in static models (Deaton (1985)) and in Collado (1998) for the dynamic model. The resulting estimator is consistent under essentially the same conditions as the estimators presented by Collado (1998). Moreover, the estimator we present is more efficient. Unfortunately, the conditions for consistency for this estimator, though relatively weak, may not be satisfied in practice.

Traditionally, the literature on repeated cross-sections is characterized by a range of alternative asymptotics (compare Verbeek (1996)). Consistency in this paper will refer to large $N$ and small $T$. That is, the number of individuals (in each cross-section) is assumed to be large, while the number of cross-sections is fixed. An additional important issue is whether the number of instruments (groups, cohorts) is kept fixed if the number of individuals increases. We follow Moffitt (1993) and treat the number of instruments as fixed, such that, asymptotically, we can ignore estimation error in the "reduced form" equations.

The remainder of this paper is organized as follows. Section 2 starts with a presentation of the model of interest and shows that OLS on the model with a predicted lagged dependent variable will generally not lead to consistent estimators. Section 3 then discusses some potentially suitable IV estimators, including those suggested by Moffitt and Collado, and argue that the necessary conditions for consistency are unlikely to be satisfied in empirical applications. Section 4 provides an illustrative example, while Section 5 concludes.

# 2 Model and Data

The observed data consists of $T$ independent cross-sections which are drawn at different points in time with each being a random sample of some underlying population. As individuals in different periods are not the same people, we index the variables by a double subscript. The $t$ denotes the cross-section and $i(t)$, $1, ..., N_t$ indexes individuals in cross-section $t$. The model is:

$$y_{i(t),t} = \alpha y_{i(t),t-1} + x'_{i(t),t}\beta + \varepsilon_{i(t),t},\ t = 1, ..., T;\ i(t) = 1, ..., N_t, \qquad (1)$$

where the $K$-dimensional vector $x_{i(t),t}$ may include time-invariant and/or time-varying variables (and a constant). We assume the error term has the following properties:

$$E\{\varepsilon_{i(t),t}x_{i(t),t}\} = 0,\ t = 1, ..., T. \qquad (2)$$

With repeated cross-section data there are generally no lagged values for $y$ observed for any individual. Following Moffitt (1993), replace $y_{i(t),t-1}$ by an estimated value. To do so, let $z_{i(t),t}$ denote a set of variables that are observed at time $t$ and $t-1$ for all individuals in cross-section $t$, including an intercept term. This requires that $z_{i(t),t}$ is time-invariant (for example, date of birth, race, gender), but time-varying variables that can be backcasted with reasonable accuracy may also be included. Below, we shall focus explicitly upon possible choices for $z_{i(t),t}$. Now, consider the orthogonal projection in cross-section $t$ of $y_{i(t),t}$ upon $z_{i(t),t}$,

$$E^*\{y_{i(t),t}|z_{i(t),t}\} = z_{i(t),t}\delta_t,\ t = 1, ..., T, \qquad (3)$$

where $E^*$ denotes the orthogonal projection (for a given $t$). This "reduced form" can be estimated consistently by regressing $y_{i(t),t}$ on $z_{i(t),t}$ to give $\hat{\delta}_t$. Following Moffitt (1993), one obtains an estimate of $y_{i(t),t-1}$ as the predicted value from this regression, substituting the appropriate $z$ values for the individuals in cross-section $t$. That is,

$$\hat{y}_{i(t),t-1} = z_{i(t),t-1}\hat{\delta}_{t-1}, \qquad (4)$$

noting that $\hat{\delta}_{t-1}$ is estimated from data on different individuals than those indexed by $i(t)$. Insert these predicted values into the original model to get:

$$y_{i(t),t} = \alpha \hat{y}_{i(t),t-1} + x'_{i(t),t}\beta + \varepsilon^*_{i(t),t}, \ t = 1, ..., T; \ i(t) = 1, ..., N_t, \qquad (5)$$

where

$$\varepsilon^*_{i(t),t} = \varepsilon_{i(t),t} + \alpha(y_{i(t),t-1} - \hat{y}_{i(t),t-1}). \qquad (6)$$

One of the explanatory variables is now measured with error, although the measurement error will be asymptotically uncorrelated with the predicted value.[1] Moffitt (1993) suggests that (5) can be consistently estimated by OLS, even if $z_{i(t),t}$ only includes an intercept term and $\hat{y}_{i(t),t-1}$ is the sample average in cross-section $t-1$. However, OLS will only produce consistent estimates of $\alpha$ and $\beta$ if (asymptotically)

$$E\{\varepsilon_{i(t),t}\hat{y}_{i(t),t-1}\} = 0, \ t = 1, ..., T \qquad (7)$$

$$E\{(y_{i(t),t-1} - \hat{y}_{i(t),t-1})x_{i(t),t}\} = 0, \ t = 1, ..., T, \qquad (8)$$

where we used (2) and the result that prediction error and predictor are orthogonal. The assumption that $\varepsilon_{i(t),t}$ is uncorrelated with $\hat{y}_{i(t),t-1}$ can be defended by the usual IV assumption that

$$E\{\varepsilon_{i(t),t}z_{i(t),t-1}\} = 0, \ t = 1, ..., T. \qquad (9)$$

However, it is unlikely that $x_{i(t),t}$ is uncorrelated with the "prediction error" $y_{i(t),t-1} - \hat{y}_{i(t),t-1}$. Consider, for example, where high $x$-values in one period on average correspond with high $x$-values in the next period. If the $\beta$ coefficients are positive this will generally imply that a high value for $x_{i(t),t-1}$, which is unobservable, will result in an underprediction of $y_{i(t),t-1}$. On the other hand, $x_{i(t),t-1}$ is positively correlated with $x_{i(t),t}$. Consequently, this will produce a positive correlation between $\varepsilon^*_{i(t),t}$ and $x_{i(t),t}$, resulting in an inconsistent estimator for $\beta$. This inconsistency carries over to $\alpha$ unless $\hat{y}_{i(t),t-1}$ is asymptotically uncorrelated with $x_{i(t),t}$. Thus this approach does not produce consistent estimates unless there are either no time-varying exogenous regressors or the time-varying exogenous variables do not exhibit any autocorrelation.

An issue we have ignored above is that $\hat{\delta}_t$ is estimated with error. Clearly, if it can be assumed that $p\lim \hat{\delta}_t = \delta_t$ (for $N_t \to \infty$) this is a small sample

[1] Unlike the standard textbook measurement error examples.

problem, as $\hat{\delta}_t$ is estimated using individuals from the previous cross-section and thus uncorrelated with any variable from cross-section $t$. This is the argument exploited in Moffitt (1993). A problem may arise, however, if the number of instruments increases with $N_t$, which occurs in the approach of Collado (1998) which we consider below.

# 3 Instrumental variables estimators

To overcome the problem of correlation between the regressors and the error term one may employ instrumental variables estimators. Note that we now need instruments for $x_{i(t),t}$ even though these variables are exogenous in the original model. Let $w_{i(t),t}$ denote an $R$-dimensional vector of potential instruments that can be used to estimate (5). Necessary conditions for consistent estimation are that (asymptotically)[2]

$$E\{\varepsilon_{i(t),t} w_{i(t),t}\} = 0, \; t = 1, ..., T \tag{10}$$

$$E\{(y_{i(t),t-1} - \hat{y}_{i(t),t-1}) w_{i(t),t}\} = 0, \; t = 1, ..., T \tag{11}$$

$$E\begin{pmatrix} w_{i(1),1}x'_{i(1),1} & w_{i(1),1}y_{i(1),0} \\ w_{i(2),2}x'_{i(2),2} & w_{i(2),2}y_{i(2),1} \\ \vdots & \vdots \\ w_{i(T),T}x'_{i(T),T} & w_{i(T),T}y_{i(T),T-1} \end{pmatrix} = \begin{pmatrix} \Sigma_1 \\ \Sigma_2 \\ \vdots \\ \Sigma_T \end{pmatrix} = \Sigma \text{ has rank } K+1. \tag{12}$$

The first condition says that the instruments should be exogenous, which is similar to condition (9). The second condition (11) implies that the prediction error is orthogonal to the instruments. That is, the instruments $w_{i(t),t}$ should not be able to predict any of the variation in $y_{i(t),t-1}$ left unexplained by $z_{i(t),t-1}$. Consequently, a natural choice is $w_{i(t),t} = z_{i(t),t-1}$, such that both (10) and (11) are automatically satisfied under condition (9). The third condition says that $TR \times (K+1)$ cross-moment matrix $\Sigma$ has full column rank. This requires that the instruments capture variation in $y_{i(t),t-1}$ independent of the variation in $x_{i(t),t}$.

It is important to note that the rank condition for identification in (12) is in terms of the true "reduced form" parameters, like $\delta_t$. This implies that

[2] We interpret these conditions as being required for each period $t$. On the pooled sample, this implies that the instruments in $w_{i(t),t}$ may be interacted with time dummies. Alternatively, the conditions can be imposed to hold over the pooled sample only. In this case, consistency is to be interpreted for $T$ (or both $T$ and $N_t$) going to infinity.

these order conditions cannot be directly verified from the data. That is, estimation error in the reduced form parameters may hide collinearity problems and thereby provide spurious identification.

In model (1), the total variation in $y_{i(t),t-1}$ can be attributed to variation in each of $x_{i(t),t-1}$, $\varepsilon_{i(t),t-1}$ and $y_{i(t),t-2}$. As a similar argument holds for $y_{i(t),t-2}$, one can say that the variation in the lagged dependent variable is due to variation in lagged error terms, lagged exogenous variables and, potentially, different starting dates or starting values. In most practical cases, the instruments will be time-invariant (e.g., date of birth), or, which is equivalent for our purposes, fully predictable from their own past (e.g., age). This implies that instruments uncorrelated with $\varepsilon_{i(t),t}$ will usually be uncorrelated with any lagged error as well. Consequently, the conditions above require that the instruments be able to predict variation in *lagged* exogenous variables different from that predicted in *current* values.[3] At the same time, they should be uncorrelated with the equation's error.

Given these stringent requirements it is likely that the number of potential instruments is small. First, the instruments (dated $t-1$) must be observed for all individuals in cross-section $t$. Although not necessarily, this will usually restrict attention to time-invariant variables. Given the conditions above and to simplify the discussion we shall now assume that $w_{i(t),t} = z_{i(t),t-1} = w_{i(t)}$, that is, the instruments used to predict $y_{i(t),t-1}$ are time-invariant and identical to those used to estimate the resulting model. In this case, (11) is automatically satisfied and conditions (9) and (10) are identical. It is well known that taking group averages is equivalent to instrumental variables estimation with the group dummies as instruments (Angrist (1991), Moffitt (1993)). Thus, when the instruments are cohort dummies, estimation of (5) by instrumental variables is identical to applying OLS to the original model where all variables are replaced by their (time-specific) cohort sample averages.

The availability of appropriate instruments, that satisfy conditions (9) and (10)-(12) may be rather limited. Thus it is natural to ask whether we can somewhat relax the requirements on the instruments. One possibility is to explicitly capture the cohort effects by including cohort fixed effects, as in Deaton (1985) and Collado (1998). This is done by including $w_{i(t)}$ as additional regressors in (5) but with *time-invariant* coefficients. Thus, we

[3] There is no need to instrument time-invariant $x$'s, provided that they are included in the instrument vector $z$.

allow for "cohort effects" by replacing assumption (10) with

$$E^*\{\varepsilon^*_{i(t),t}|w_{i(t)}\} = w'_{i(t)}\lambda. \tag{13}$$

If the instruments are correlated with the error term, it might be expected that (7) is violated and that the predicted value $\hat{y}_{i(t),t-1}$ is correlated with the equation's error. If, however, this correlation is time-invariant in the sense that

$$E^*\{\varepsilon^*_{i(t),t}|\hat{y}_{i(t),t-1}\} = w'_{i(t)}\lambda_2 \tag{14}$$

this is automatically taken into account with (13).[4] Thus we can write

$$y_{i(t),t} = \alpha\hat{y}_{i(t),t-1} + x'_{i(t),t}\beta + w'_{i(t)}\lambda + \eta_{i(t),t}, \tag{15}$$

where

$$E\{\eta_{i(t),t}w_{i(t)}\} = 0,\ t = 1, ..., T \tag{16}$$

replaces condition (10). This also allows one to relax assumption (2) to

$$E\{\eta_{i(t),t}x_{i(t),t}\} = 0,\ t = 1, ..., T. \tag{17}$$

Thus we would estimate (15) by IV using $w_{i(t)}$, interacted with time dummies, as instruments. We shall refer to this as the *augmented IV estimator* noting that a time-varying $\lambda$ would make the model unidentified. To achieve identification, we need to strengthen (12) to

$$\begin{pmatrix} \Sigma_2 - \Sigma_1 \\ \vdots \\ \Sigma_T - \Sigma_1 \end{pmatrix} \text{ has rank } K + 1. \tag{18}$$

If $\Sigma_t$ does not vary sufficiently over the cross-sections, the variation of the instruments is collinear with $w_{i(t)}$ and estimation will break down. A direct implication of this is that one needs at least three cross-sections ($t = 0, 1, 2$) to identify the model under these assumptions.

Computation of this augmented IV estimator is remarkably simple if $w_{i(t)}$ is a set of cohort dummies. First, one simply aggregates the data into cohort averages, which gives

$$\hat{y}_{i(t),t} = \alpha\hat{y}_{i(t),t-1} + \hat{x}'_{i(t),t}\beta + w'_{i(t)}\lambda + \hat{\eta}_{i(t),t}, \tag{19}$$

[4] Here, our assumption that $\hat{\delta}_t$ converges to $\delta_t$ is important. That is, asymptotically knowing $w_{i(t)}$ and $\hat{y}_{i(t),t-1}$ is equivalent to knowing $w_{i(t)}$.

where the hats denote predicted values from a period-by-period regression on $w_{i(t)}$, that is, they denote (sample) cohort averages. Second, applying OLS to (19) corresponds to the standard within estimator for $(\alpha, \beta')'$ based upon treating the cohort-level data as a panel. The usual problem with estimating dynamic panel data models (see Nickell (1981)), does not arise here because under assumption (16) the error term, which is a within cohort average of individual error terms that are uncorrelated with $w_{i(t)}$, is asymptotically zero.[5] However, it remains whether suitable instruments can be found that satisfy the above conditions.

We are now in a position to present Collado's (1998) estimator.[6] As Collado explicitly restricts attention to models in cohort averages, the instruments $w_{i(t)}$ are time-invariant and to be interpreted as a set of cohort dummies. The main extension is that condition (16) is relaxed. In particular, a *time-varying* correlation between $\hat{y}_{i(t),t-1}$ and $\eta_{i(t),t}$ is now allowed. Following standard dynamic panel data IV procedures, the cohort fixed effects $\lambda$ are first eliminated by first-differencing, to give

$$\begin{aligned}\hat{y}_{i(t),t} - \hat{y}_{i(t-1),t-1} &= \alpha(\hat{y}_{i(t),t-1} - \hat{y}_{i(t-1),t-2}) \\ &\quad +(\hat{x}_{i(t),t} - \hat{x}_{i(t-1),t-1})'\beta + (\hat{\eta}_{i(t),t} - \hat{\eta}_{i(t-1),t-1}),\end{aligned} \tag{20}$$

where it can be noted that, because of the definition of $w_{i(t),t}$ as cohort dummies, $\hat{y}_{i(t-1),t-1} = \hat{y}_{i(t),t-1}$ and similarly for the other variables and lags.[7] In other words, in (20) we have a standard (first-differenced) dynamic panel data model with the unit of observation being a cohort. If condition (16) is not valid, the cohort averages $\hat{\eta}_{i(t),t}$ are asymptotically nonzero and correlated with $w_{i(t)}$ in a time-varying way. Furthermore, $\hat{y}_{i(t),t-1}$ and $\hat{\eta}_{i(t-1),t-1}$ are correlated by construction, and OLS applied to (20) is inconsistent. Consequently, equation (20) must also be estimated by IV. For the genuine panel data model, Arellano and Bond (1991) provide a list of instruments which is exploited by Collado (1998). The essential condition for identification, however, is that $\hat{y}_{i(t-1),t-2}$, or one of its lags, is a valid instrument for $\hat{y}_{i(t),t-1} - \hat{y}_{i(t-1),t-2}$. This, as before, requires time variation in $\hat{y}_{i(t),t-1}$, but also requires that

$$E\{(\hat{\eta}_{i(t),t} - \hat{\eta}_{i(t-1),t-1})\hat{y}_{i(t-1),t-2}\} = 0, \tag{21}$$

[5]Recall that, asymptotically, the number of cohorts is fixed and the number of individuals goes to infinity.

[6]We shall restrict attention to the estimator that is claimed to be consistent for fixed $T$.

[7]Collado (1998) writes $y_{c,t}$, where $c$ indexes cohorts.

which is trivially satisfied if there are no cohort effects in $\eta_{i(t),t}$ (as with the augmented IV estimator).[8] In general, however, it is less obvious that this condition is satisfied.[9] If

$$E^*\{\eta_{i(t),t}|w_{i(t)}\} = w'_{i(t)}\phi_t,\ t = 1, ..., T \tag{22}$$

we thus need that (for finite $T$)

$$E(w'_{i(t)}(\phi_t - \phi_{t-1}) \cdot w'_{i(t)}\delta_{t-2}\} = 0, \tag{23}$$

which, given the definition of $w_{i(t)}$, requires that $(\phi_t - \phi_{t-1})'\delta_{t-2} = 0$. Collado (1998) imposes this condition while assuming that the dimension of $w_{i(t)}$ (and thus the dimensions of $\phi_t$ and $\delta_t$) increases with sample size, through imposing that $\phi_t$ is orthogonal to $\delta_s$, $s \neq t$. In our case, when we keep the dimension of $w_{i(t)}$ fixed, the only case where the condition can be expected to be satisfied is where $\phi$ is time-invariant and thus included in $\lambda$. The conditions for consistency of this estimator are thus no weaker than those for the augmented IV estimator.

A different way of stating this last result is that there is no gain from instrumenting $\hat{y}_{i(t),t-1} - \hat{y}_{i(t-1),t-2}$ in (20). On the contrary, unnecessarily doing so will lead to a loss in efficiency compared to the augmented IV estimator, while the gain in robustness is spurious. As a result, estimating a dynamic model from cohort level data can be much simpler than from genuine panel data. The reason is that consistency requires that at the level of cohort aggregates, there is no longer a time-specific correlation between lagged dependent variable and lagged error term. On the other hand, if such correlation exists, any further lag of the dependent variable would not provide a valid instrument.

[8] Note that any time-invariant cohort effects would have been eliminated by the first-difference transformation.

[9] Recall that we ignore the measurement error problem in Collado (1998) by assuming that $\hat{\delta}_t$, and the other "reduced form" parameters, asymptotically converge to their true values. In Collado's approach (following Deaton (1985)) the number of cohorts is assumed to grow with sample size. This implies that the number of instruments is increasing with $N_t$ and it is no longer obvious that $\hat{\delta}_t - \delta_t$ is asymptotically zero (note that the definition of $\delta_t$ changes with sample size). Keeping the number of instruments fixed, as we shall do, the estimation error in $\hat{\delta}_t$ is a small sample problem, which may or may not be negligible (see Verbeek and Nijman (1992) for a discussion of this issue in the static model). Once we have a consistent estimator for $\alpha$ and $\beta$ this can be addressed.

# 4 A simple example

To illustrate and clarify the conditions for consistency of the respective estimators, as discussed above, consider a simple example where the model of interest contains only one exogenous variable. The data generating process is given by

$$y_{it} = \alpha y_{i,t-1} + \beta x_{it} + \varepsilon_{it}, \qquad 0 < \alpha < 1, \tag{24}$$

where the error term has the usual error components structure

$$\varepsilon_{it} = \theta_i + v_{it}. \tag{25}$$

It is assumed that $v_{it}$ is uncorrelated over time. Let $w_i$ denote a set of time-invariant variables that are used as instruments. The main question is what conditions need to be imposed upon the relationships between $w_i$ and the other variables in the model to guarantee that one or more of the estimators are consistent.

The first step, for each estimator, is to approximate the lagged dependent variable by a linear projection upon $w_i$, denoted $E^*\{y_{i,t-1}|w_i\}$. Using recursive substitution, one can easily derive that

$$\begin{aligned} E^*\{y_{i,t-1}|w_i\} &= \beta \sum_{j=0}^{t-2} \alpha^j E^*\{x_{i,t-j-1}|w_i\} + \left(\sum_{j=0}^{t-2} \alpha^j\right) E^*\{\theta_i|w_i\} \\ &\quad + \alpha^{t-1} E^*\{y_{io}|w_i\} + \sum_{j=0}^{t-2} \alpha^j E^*\{v_{i,t-j-1}|w_i\}. \end{aligned} \tag{26}$$

In a first case, case A, we assume that the starting value $y_{io}$ is exogenous and independent of $\theta_i$.[10] In a second case, case B, we assume that the process is in equilibrium or - equivalently - that it started in an infinite past. For case B, the expression for $E^*\{y_{i,t-1}|w_i\}$ simplifies to

$$\begin{aligned} E^*\{y_{i,t-1}|w_i\} &= \beta \sum_{j=0}^{\infty} \alpha^j E^*\{x_{i,t-j-1}|w_i\} \\ &\quad + \frac{1}{1-\alpha} E^*\{\theta_i|w_i\} + \sum_{j=0}^{\infty} \alpha^j E^*\{v_{i,t-j-1}|w_i\}. \end{aligned} \tag{27}$$

Note that in the first case the instruments $w_i$ may be predicting the lagged dependent variable partly through the initial value $y_{io}$.

[10] The starting date need not coincide with the beginning of the sample period.

Let us consider the respective estimators. First, the OLS estimator requires that the instruments are uncorrelated with the equation's error terms. That is

$$E^*\{\theta_i|w_i\} = 0 \tag{28}$$
$$E^*\{v_{it}|w_i\} = 0. \tag{29}$$

Further, it is required that the prediction error is uncorrelated with $x_{it}$, which requires that $x_{i,t-j-1} - E^*\{x_{i,t-j-1}|w_i\}$ is uncorrelated with $x_{it}$ $(j = 0, 1, 2, ..)$. For a time-varying $x_{it}$-variable, this imposes the strong restriction of the absence of autocorrelation relative to the cohort-specific mean.

The standard IV estimator, using $w_i$ as instruments for $x_{it}$ as well, also imposes conditions (28)-(29). In addition, the rank condition requires that $E^*\{x_{it}|w_i\}$ is not collinear with $E^*\{y_{i,t-1}|w_i\}$. For case A it is sufficient that $E^*\{x_{it}|w_i\} \neq 0$ (and $\beta \neq 0$) or that $E^*\{y_{io}|w_i\} \neq 0$ (and $\alpha \neq 0$). For case B we need that $E^*\{x_{it}|w_i\}$ varies with $t$. For case B it is thus required that the time-invariant variables in $w_i$ have a time-varying relationship with the exogenous variables in the model. For case A this is not required as the fixed starting date of the process induces variation over time even if the relationship between $w_i$ and $x_{it}$ is time-invariant.[11]

For the augmented IV estimator, which extends the standard IV estimator by including $w_i$ in the model, condition (28) is no longer required. Implicitly, this allows for cohort-specific means in the processes for $y_{it}$ and $x_{it}$. Both case A and case B now require that $E^*\{x_{it}|w_i\}$ varies with $t$ (and $\beta \neq 0$). For case A, this may again be replaced by $E^*\{y_{io}|w_i\} \neq 0$. It should be stressed that, given the inclusion of $w_i$ in the model, the need for time-variation in $E^*\{x_{it}|w_i\}$ and $E^*\{y_{i,t-1}|w_i\}$ is much stronger.

Finally, consider a simple variant of the estimator proposed by Collado (1998), which restricts attention to discrete-valued $w_i$. While this estimator does not require conditions (28)-(29), it does require that

$$E^*\{v_{it}|w_i\} - E^*\{v_{i,t-1}|w_i\}$$

is uncorrelated with $E^*\{v_{i,t-j}|w_i\}$ for $j = 2, 3, ....$ Note that the case with a time-invariant $E^*\{v_{it}|w_i\}$ is already captured by the augmented IV estimator through the inclusion of $w_i$. The only relevant extension thus occurs with a

[11] If $\alpha$ is small or $t$ is large, the impact of the initial values is small and time-variation in $E^*\{x_{it}|w_i\}$ is recommended (as in case B).

time-varying $E^*\{v_{it}|w_i\}$. However, the only stationary process that is consistent with this condition is characterized by equi-correlation and implies that $E^*\{v_{it}|w_i\}$ is zero (given the presence of the cohort-specific effect $E\{\theta_i|w_i\}$). As a result, the relaxed conditions seem to be no weaker than those for the augmented IV estimator and the additional stage of instrumentation in Collado (1998) appears unnecessary.

# 5 Concluding remarks

Moffitt (1993) and Collado (1998) consider the estimation of dynamic models on the basis of repeated cross-sections to be a relatively straightforward extension of the static fixed effects model (see, for example, Deaton (1985)). This paper shows that the extension is nontrivial and severely restricts the set of suitable instruments. Moffitt's estimator is shown to be inconsistent, unless the exogenous variables are either time-invariant or exhibit no autocorrelation. Alternative instrumental variables estimators are consistent for the model parameters only in restrictive cases. We propose an augmented IV estimator, corresponding to the standard within estimator applied to the pseudo panel of cohort averages, that is consistent under essentially the same conditions as the more complicated estimators suggested by Collado (1998). Furthermore, our proposed estimator is more efficient. However, necessary for these estimators is the requirement that the time-invariant instruments have a time-varying relationship with each of the exogenous variables in the model. While it is not obvious that this requirement will be satisfied in empirical applications, it is also not easy to check, because estimation error in the reduced form parameters may hide collinearity problems.

# References


[1] Anderson, T.W. and C. Hsiao (1981), "Estimation of Dynamic Models with Error Components", *Journal of the American Statistical Association*, 76, 598-606.

[2] Angrist, J.D. (1991), "Grouped-data estimation and testing in simple labor-supply models", *Journal of Econometrics*, 47, 243-266.

[3] Arellano, M. and S. Bond (1991), "Some Test of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations", *Review of Economic Studies*, 58, 277-297.

[4] Collado, M.D. (1998), "Estimating Dynamic Models from Time Series of Independent Cross-Sections", *Journal of Econometrics,* 82, 37-62.

[5] Deaton, A. (1985), "Panel Data from Times Series of Cross-Sections", *Journal of Econometrics*, 30, 109-126.

[6] Moffitt, R. (1993), "Identification and Estimation of Dynamic Models with a Time Series of Repeated Cross-Sections", *Journal of Econometrics*, 59, 99-124.

[7] Nickell, S. (1981), "Biases in Dynamic Models with Fixed Effects", *Econometrica*, 49, 1417-1426.

[8] Verbeek, M. (1996), "Pseudo Panel Data", in: L. Mátyás and P. Sevestre, eds., *The Econometrics of Panel Data: Handbook of the Theory with Applications*, 2nd revised edition, Kluwer Academic Publishers, 280-292.

[9] Verbeek, M. and Th. Nijman (1992), "Can Cohort Data Be Treated as Genuine Panel Data?", *Empirical Economics*, 17, 9-23.

[10] Verbeek, M. and Th. Nijman (1993),"Minimum MSE Estimation of a Regression Model with Fixed Effects from a Series of Cross Sections", *Journal of Econometrics*, 59, 125-136.