MOMENT CONDITIONS FOR AR(1) PANEL DATA MODELS WITH MISSING OUTCOMES

David Pacini
Frank Windmeijer

Discussion Paper 15 / 660

26 May 2015

Department of Economics
University of Bristol
8 Woodland Road
Bristol BS8 1TN
United Kingdom

Moment Conditions for AR(1) Panel Data Models with Missing Outcomes*

David Pacini and Frank Windmeijer
Dept of Economics, University of Bristol
david.pacini@bristol.ac.uk; f.windmeijer@bristol.ac.uk

May 2015

Abstract

We derive moment conditions for dynamic, AR(1) panel data models when values of the outcome variable are missing. In this context, commonly used estimators only use data on individuals observed for at least three consecutive periods. We derive moment conditions for observations with at least three non-consecutive observations for estimation of the parameters by GMM.
Keywords: Panel Data, Missing Values.
JEL Classification: C33, C51

1 Introduction

We derive moment conditions for the AR(1) dynamic panel data model when values of the outcome variable are missing completely at random. When data on the outcome variable are available for all periods for all cross-sectional units (referred to as individuals henceforth) the parameters of interest in linear dynamic panel data models can be estimated efficiently using a generalised method-of-moment (GMM) estimator based on the moment restrictions described in e.g. Ahn and Schmidt (1995, 1997).

Denoting T the number of panel time periods, under standard assumptions there are $(T-1)(T-2)/2$ linear and $(T-3)$ nonlinear moment conditions available, needing at least 3 and 4 consecutive observations respectively. In practice, most commonly used estimation procedures discard the information from the individuals that do not have the required number of consecutive observations, or only use that part of the observations

*We would like to thank Pedro Albarran, Jon Temple and an anonymous referee for helpful comments.

that do. For instance, in a $T = 8$ panel, for those individuals with missing outcomes in periods 5 and 8, only the consecutive observations in periods 1 to 4 contribute 4 moment conditions for the estimation of the parameters.

In contrast, if we denote by T_i the number of observed outcomes for individual i, we derive $(T_i - 1)(T_i - 2)/2 + (T_i - 3)$ moment conditions for any $T_i \geq 3$, allowing for any missing data pattern. They can be represented in terms of first differences in the observed observations, but with varying parameters. These are simple functions of the model parameters and the number of periods between successive observed outcomes.

Several previous studies consider estimation of dynamic panel data models with outcomes either missing-at-random (e.g., Arellano and Bond, 1991; Abrevaya, 2013; Albarran and Arellano, 2013) or missing not-at-random (e.g., Arellano, Bover and Labeaga, 1997; Semykina and Wooldridge, 2013). Most methods have focused on the available linear moment conditions for missing data patterns that have at least three consecutive observations, For example, the commonly used "pooled GMM" estimator uses information from at least 3 consecutive observations, but sets the value of any missing observation to 0 when it is used as an instrument. The "expanded GMM" treats this latter situation as a different reduced form. This is essentially the method proposed by Muris (2013), who estimates the parameters by minimum distance after estimation on all subsamples with at least three consecutive observations separately. Albarran and Arellano (2013) propose a "cross-sample GMM" estimator that utilises information that effectively reduces the number of estimated reduced form parameters. None of these methods use any of the information contained in non-consecutive observations.

We consider the case when outcomes are missing-at-random, but other explanatory variables are always observed. This is a situation akin to Kniesner et al. (2012), where the outcomes, individual wages, were missing for some individuals in some periods, but the industry level accident rate as the explanatory variable to determine the value of life was always observed.

We first derive in Section 3 the moment conditions for panels with missing cross-sections, i.e. outcomes are missing for all individuals in some periods. In this situation, no additional assumptions are needed for the derived moment conditions to be valid. In Section 4 we generalise the findings to the case where different individuals can have different missing data patterns and state the assumptions needed for the moment conditions to remain valid. Both sections present some Monte Carlo results for a GMM estimator utilising these new moment conditions.

2 Moments for the AR(1) Panel Data Model

We consider the standard AR(1) panel data model of the form

$$\begin{aligned} y_{it} &= \alpha y_{i,t-1} + u_{it} \\ u_{it} &= \eta_i + v_{it} \end{aligned} \tag{1}$$

for $i = 1, ..., n$ and $t = 2, ..., T$. n is large, T is fixed and $|\alpha| < 1$. η_i and v_{it} have the familiar error component structure with assumptions (for all i and t),

$$E(\eta_i) = E(v_{it}) = E(\eta_i v_{it}) = 0, \tag{2}$$

$$E(v_{it} v_{is}) = 0, \forall t \neq s. \tag{3}$$

In addition, there is the standard assumption concerning the initial condition y_{i1}

$$E(y_{i1} v_{it}) = 0, \tag{4}$$

for $i = 1, ..., n$ and $t = 2, ..., T$. Assumptions (2), (3) and (4) are the standard assumptions implying moment conditions that are sufficient to identify and estimate α for $T \geq 3$.

Arellano and Bond (1991) use the set of $(T-1)(T-2)/2$ available linear moment conditions

$$E\left(y_i^{t-2}(u_{it} - u_{it-1})\right) = E\left(y_i^{t-2}(\Delta y_{it} - \alpha \Delta y_{i,t-1})\right) = 0, \tag{5}$$

for $t = 3, ..., T$, where $y_i^{t-2} = (y_{i1}, y_{i2}, ..., y_{i,t-2})'$, and $\Delta y_{it} = y_{it} - y_{i,t-1}$. For $t = 4, ..., T$, Ahn and Schmidt (1995) show that there are $(T-3)$ additional non-linear moment conditions available under the standard assumptions:

$$E(u_{it} \Delta u_{i,t-1}) = E((y_{it} - \alpha y_{i,t-1})(\Delta y_{i,t-1} - \alpha \Delta y_{i,t-2})) = 0. \tag{6}$$

3 Missing Cross Sections

Consider first the situation where outcomes in some periods are not observed for any individual, i.e. there are gaps in the panel survey. Let y_i^o be the vector of observed outcomes for individual i, and denote by $T_i = T^o$ the number of observations in y_i^o. Denote by $t(j)$ the panel time period of the j-th observation $y_{ij}^o = y_{i,t(j)}$. Let $d_k = t(k) - t(k-1)$ be the number of periods between y_{ik}^o and $y_{i,k-1}^o$. For example, for $T = 6$, with missing observations in periods 3 and 4, $y_i^o = (y_{i1}, y_{i2}, y_{i5}, y_{i6})'$, $T^o = 4$, $(d_2, d_3, d_4) = (1, 3, 1)$ and $\{t(1), t(2), t(3), t(4)\} = \{1, 2, 5, 6\}$.

From the model structure (1) it follows that $y_{it} = \alpha^p y_{i,t-p} + \eta_i \sum_{r=0}^{p-1} \alpha^r + \sum_{l=0}^{p-1} \alpha^l v_{i,t-l}$, and hence $y_{ij}^o = \alpha^{d_j} y_{i,j-1}^o + \eta_i \sum_{r=0}^{d_j-1} \alpha^r + \sum_{l=0}^{d_j-1} \alpha^l v_{i,t(j)-l}$. We can therefore write the model for the observed outcomes as

$$A(\alpha) y_i^o = \tau(\alpha) \eta_i + \widetilde{v}_i(\alpha), \tag{7}$$

where $A(\alpha)$ and $\tau(\alpha)$ are a $(T^o - 1) \times T^o$ matrix and a $(T^o - 1)$ vector respectively, defined as

$$A(\alpha) = \begin{pmatrix} -\alpha^{d_2} & 1 & 0 & \cdots & 0 \\ 0 & -\alpha^{d_3} & 1 & & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \alpha^{-d_{T^o}} & 1 \end{pmatrix}; \ \tau(\alpha) = \begin{pmatrix} \sum_{r=0}^{d_2-1} \alpha^r \\ \sum_{r=0}^{d_3-1} \alpha^r \\ \vdots \\ \sum_{r=0}^{d_{T^o}-1} \alpha^r \end{pmatrix} = \frac{1}{1-a} \begin{pmatrix} 1-\alpha^{d_2} \\ 1-\alpha^{d_3} \\ \vdots \\ 1-\alpha^{d_{T^o}} \end{pmatrix}$$

and

$$\widetilde{v}_i(\alpha) = \begin{pmatrix} \sum_{l=0}^{d_2-1} \alpha^l v_{i,t(2)-l} \\ \sum_{l=0}^{d_3-1} \alpha^l v_{i,t(3)-l} \\ \vdots \\ \sum_{l=0}^{d_{T^o}-1} \alpha^l v_{i,t(T^o)-l} \end{pmatrix}.$$

As in (5), we need to find a transformation that eliminate η_i and is such that past observations are not correlated with the transformed errors. A transformation matrix that satisfies these conditions is given by

$$G(a) = \begin{pmatrix} -\frac{1-\alpha^{d_3}}{1-\alpha^{d_2}} & 1 & 0 & \cdots & 0 \\ 0 & -\frac{1-\alpha^{d_4}}{1-\alpha^{d_3}} & 1 & & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & -\frac{1-\alpha^{d_{T^o}}}{1-\alpha^{d_{T^o-1}}} & 1 \end{pmatrix},$$

as then $G(\alpha)\tau(\alpha) = 0$, and hence

$$G(\alpha) A(\alpha) y_i^o = G(\alpha) \widetilde{v}_i(\alpha) \equiv \widetilde{v}_i^*,$$

with

$$\widetilde{v}_i^* = \begin{pmatrix} \sum_{l=0}^{d_3-1} \alpha^l v_{i,t(3)-l} - \frac{1-\alpha^{d_3}}{1-\alpha^{d_2}} \left(\sum_{l=0}^{d_2-1} \alpha^l v_{i,t(2)-l} \right) \\ \sum_{l=0}^{d_4-1} \alpha^l v_{i,t(4)-l} - \frac{1-\alpha^{d_4}}{1-\alpha^{d_3}} \left(\sum_{l=0}^{d_3-1} \alpha^l v_{i,t(3)-l} \right) \\ \vdots \\ \left(\sum_{l=0}^{d_{T^o}-1} \alpha^l v_{i,t(T^o)-l} \right) - \frac{1-\alpha^{d_{T^o}}}{1-\alpha^{d_{T^o-1}}} \left(\sum_{l=0}^{d_{T^o-1}-1} \alpha^l v_{i,t(T^o-1)-l} \right) \end{pmatrix}. \tag{8}$$

The earliest idiosyncratic shock appearing in the first element of $\widetilde{v}_i^*$ is $v_{i,t(1)+1}$, and hence $E\left[y_{i,t(1)}\widetilde{v}_{i1}^*\right] = E\left[y_{i1}^o \widetilde{v}_{i1}^*\right] = 0$. Similarly, the earliest idiosyncratic shock appearing in $\widetilde{v}_{i2}^*$ is $v_{i,t(2)+1}$, resulting in $E\left[y_{i,t(1)}\widetilde{v}_{i2}^*\right] = E\left[y_{i,t(2)}\widetilde{v}_{i2}^*\right] = E\left[y_{i1}^o \widetilde{v}_{i2}^*\right] = E\left[y_{i2}^o \widetilde{v}_{i2}^*\right] 0$, and so forth. We therefore obtain the following $(T^o-1)(T^o-2)/2$ moment conditions

$$E\left[Z_i^{o\prime} G(\alpha) A(\alpha) y_i^o\right] = 0,$$

where

$$Z_i^o = \begin{pmatrix} y_{i1}^o & 0 & 0 & \cdots & \cdots & \cdots & \cdots & 0 \\ 0 & y_{i1}^o & y_{i2}^o & & & & & \vdots \\ \vdots & & & \ddots & & & & 0 \\ 0 & \cdots & \cdots & 0 & y_{i1}^o & y_{i2}^o & \cdots & y_{i,T^o-2}^o \end{pmatrix}.$$

The transformation $G(\alpha) A(\alpha) y_i^o$ can be simplified. For $j = 3, ..., T^o$, the $(j-2)$-th element of $G(\alpha) A(\alpha) y_i^o$ is given by

$$\begin{aligned} & \left(y_{ij}^o - \alpha^{d_j} y_{i,j-1}^o\right) - \frac{1-\alpha^{d_j}}{1-\alpha^{d_{j-1}}}\left(y_{i,j-1}^o - \alpha^{d_{j-1}} y_{i,j-2}^o\right) \\ = \; & \left(y_{ij}^o - y_{i,j-1}^o\right) - \alpha^{d_{j-1}} \frac{1-\alpha^{d_j}}{1-\alpha^{d_{j-1}}}\left(y_{i,j-1}^o - y_{i,j-2}^o\right) \\ = \; & \Delta y_{ij}^o - \alpha^{d_{j-1}} \frac{1-\alpha^{d_j}}{1-\alpha^{d_{j-1}}} \Delta y_{i,j-1}^o, \end{aligned}$$

resulting in the $(T^o-1)(T^o-2)/2$ moment conditions

$$E\left[\left(y_{i1}^o, ..., y_{i,j-2}^o\right)' \left(\Delta y_{ij}^o - \alpha^{d_{j-1}} \frac{1-\alpha^{d_j}}{1-\alpha^{d_{j-1}}} \Delta y_{i,j-1}^o\right)\right] = 0, \tag{9}$$

for $j = 3, ..., T^o$. These moment conditions are based on simple first differences in any observed sequence of outcomes, with differing parameters which are simple functions of α and the number of periods between observed outcomes only.

We can further derive $T^o - 3$ moment conditions equivalent to the Ahn-Schmidt nonlinear moment conditions (6). [1]From the results derived above, we get that, for

[1]We refrain here from making further assumptions about the initial condition y_{i1}. However, it is easily established that if the Arellano and Bover (1995) and Blundell and Bond (1998) moments $E\left[\Delta y_{i,t-1}\left(y_{it} - \alpha y_{i,t-1}\right)\right] = 0$, for $t = 3, ..., T$ hold in the complete panel, then $E\left[\Delta y_{i,j-1}^o\left(y_{ij}^o - \alpha^{d_j} y_{i,j-1}^o\right)\right] = 0$ for $j = 3, ..., T^o$.

$j = 4, .., T^o$,

$$E\left[\left(y^o_{ij} - \alpha^{d_j} y^o_{i,j-1}\right)\left(\Delta y^o_{i,j-1} - \alpha^{d_{j-2}} \frac{1-\alpha^{d_{j-1}}}{1-\alpha^{d_{j-2}}} \Delta y^o_{i,j-2}\right)\right] = 0. \tag{10}$$

3.1 Estimation

We present some Monte Carlo estimation results for a $T = 6$ panel with various combinations of 2 missing cross sections. The moment conditions are nonlinear in α, and the corresponding errors are also functions of α. A two-step GMM procedure will be efficient given the moment conditions, but an initial efficient weight matrix that does not depend on estimated parameters is clearly not available. To avoid dependence on an initial weight matrix, we estimate the parameters using the Continuous Updating Estimator (CUE). Let $g_i(\alpha)$ be the vector of moment conditions for individual i. Define $\widehat{g}(\alpha) = n^{-1} \sum_{i=1}^{n} g_i(\alpha)$ and $\widehat{\Omega}(\alpha) = n^{-1} \sum_{i=1}^{n} g_i(\alpha) g_i(\alpha)'$, then the CUE is defined as

$$\widehat{\alpha}_{CUE} = \arg\min_{a} \widehat{g}(a)' \left(\widehat{\Omega}(a)\right)^{-1} \widehat{g}(a).$$

The CUE has the further advantage that it is invariant to reparameterisation of the moments, for example, it is invariant to whether we specify the moments as in (9), or equivalently as

$$E\left[\left(y^o_{i1}, ..., y^o_{i,j-2}\right)' \left(\left(1-\alpha^{d_{j-1}}\right) \Delta y^o_{ij} - \alpha^{d_{j-1}} \left(1-\alpha^{d_j}\right) \Delta y^o_{i,j-1}\right)\right] = 0.$$

It is also well behaved under many weak instrument asymptotics, see Newey and Windmeijer (2009).

Table 1 presents some Monte Carlo results for a design with $\eta_i \sim N(0,1)$; $v_{it} \sim N(0,1)$ and with the initial observation y_{i1} drawn from the covariance-stationary distribution. We set the value of α to 0.4 or 0.8. The number of Monte Carlo replications is 1000. We present Bias, Root mean squared error (Rmse), Median Bias and Interquartile Range (IQR) of the CUE estimates for samples of size $n = 1000$, and Rmse and IQR for samples of size $10,000$. Median Bias and IQR are generally more reliable measures for the CUE as some outlying CUE estimates can be obtained in any set of replications.[2]

[2] For comparison, the bias, Rmse, Med Bias and IQR of the standard two-step Arellano-Bond GMM estimator for the $\alpha = 0.8$, $n = 1000$ and no missing data case are -0.0269, 0.0897, -0.0276 and 0.1185 respectively.

For $n = 1000$, when $\alpha = 0.4$ there is virtually no bias in the estimator for any of the designs. Rmse and IQR are smallest for the case when y_3 and y_4 are missing. The Rmse and IQR are then only 32% and 29% larger than those for the case where there are no missing data. In contrast, the Rmse and IQR for the case where y_5 and y_6 are missing, i.e. we have a standard $T = 4$ panel, are about double those of the full $T = 6$ panel. This result is even stronger for the $\alpha = 0.8$ case, where the standard $T = 4$ panel case is performing the worst on all measures by a considerable amount. The best results in terms of bias, Rmse, median bias and IQR are obtained for $\alpha = 0.8$ and $n = 1000$ when y_2 and y_5 are missing. This missing data pattern actually resulted in the worst performance overall when $\alpha = 0.4$, as also confirmed for the $n = 10,000$ case. In the larger sample, the missing y_3 and y_4 case also performs well, and best on IQR, when $\alpha = 0.8$. For this missing data pattern, there are no standard moment conditions available.[3]

Table 1. Monte Carlo Estimation Results, $T = 6$

		$n = 1000$				$n = 10,000$	
	Missing obs	Bias	Rmse	Med Bias	IQR	Rmse	IQR
$\alpha = 0.4$	None	0.0020	0.0234	0.0011	0.0325	0.0072	0.0093
	y_5, y_6	0.0027	0.0481	0.0003	0.0629	0.0148	0.0198
	y_4, y_5	0.0010	0.0452	0.0007	0.0622	0.0143	0.0196
	y_3, y_5	0.0010	0.0392	0.0004	0.0530	0.0124	0.0169
	y_2, y_5	-0.0006	0.0479	0.0023	0.0639	0.0156	0.0206
	y_3, y_4	0.0017	0.0308	0.0018	0.0420	0.0098	0.0131
$\alpha = 0.8$	None	0.0240	0.0793	0.0047	0.0774	0.0149	0.0196
	y_5, y_6	0.0648	0.1697	0.0272	0.2297	0.0522	0.0575
	y_4, y_5	0.0235	0.1147	0.0015	0.1256	0.0269	0.0360
	y_3, y_5	0.0306	0.1310	0.0130	0.1148	0.0238	0.0272
	y_2, y_5	0.0185	0.1014	-0.0027	0.1053	0.0233	0.0301
	y_3, y_4	0.0471	0.1107	0.0191	0.1301	0.0245	0.0260

4 Missing Completely at Random

For the missing cross-sections case, the derived moment conditions lead to a consistent estimator of α under no further assumptions than (2)-(4), as this type of missingness is clearly exogenous and hence ignorable. In practice, we observe different missingness

[3] One could clearly specify the linear moment condition $E\left[y_{i1}\left(\left(y_{i6}-y_{i2}\right)-\alpha\left(y_{i5}-y_{i1}\right)\right)\right]=0$. This can be shown to be a linear combination of the moments (9).

patterns, and in order for the moment conditions to remain valid, we make the assumption that there is no systematic selection in the missingness generating process and that outcomes are missing completely at random.

Let s_{it} be an indicator variable equal to 1 if y_{it} is observed and zero otherwise and define $T_i = \sum_{t=1}^{T} s_{it}$. We further define the vectors $s_i = (s_{i1}, .., s_{iT})$ and $v_i = (v_{i2}, ...v_{iT})'$. We assume

Assumption M (Data are Missing Completely at Random) *Data are available from an unbalanced panel* $\{s_{it}y_{it}, s_{it}\}_{i=1,t=1}^{n,T}$ *with* $T_i \geq 3$. *For all* i *and for every possible realization* $s_r \in \mathcal{S}$ *of* s_i, *the probability* $P(s_i = s_r | y_{i1}, \eta_i, v_i) = \pi_r$, $0 < \pi_r < 1$ *does not depend on* (y_{i1}, η_i, v_i).

We assume that $T_i \geq 3$ for all i since individuals with $T_i \leq 2$ do not provide information about α. For any individual with $T_i \geq 3$ observations, the results from Section 3 produce a total of $(T_i - 1)(T_i - 2)/2 + (T_i - 3)$ moment conditions. Under Assumption M, we can combine the sets of moment conditions for groups with different missing outcome data patterns for consistent estimation of α, and improve efficiency by adding these new moment conditions.

As an illustration, consider a $T = 5$ panel with n_c individuals with a complete set of observations, $s_i = \{1, 1, 1, 1, 1\}$. For this group we have the 8 standard linear and nonlinear moment conditions. There are n_3 individuals that miss observation y_{i3}, $s_i = \{1, 1, 0, 1, 1\}$. As this latter group does not have three consecutive observations, they get omitted by most estimation procedures, (although again, the linear moment $E[y_{i1}((y_{i5} - y_{i2}) - \alpha(y_{i4} - y_{i1}))] = 0$ is available). We assume that n_c/n and n_3/n converge to constants for $n \to \infty$, and $n = n_c + n_3$. We estimate α by the CUE. As there is no overlap in the two sets of moment conditions and we assume random sampling, we get a simple block-diagonal structure.

Table 2 presents the estimation results for the same design as in Section 3.1, for $\alpha = 0.4$ and $\alpha = 0.8$, $n = 1000$ and for 1000 MC replications. The proportions n_3/n considered are 0, 0.2, 0.5, 0.8 and 1. "CC" denotes the estimator based on the standard moment conditions using only the group of individuals with a complete set of observations. "All" denotes the estimator based on all available moment conditions. For $\alpha = 0.4$, we see that the Rmse and IQR increase for "CC" with increasing n_3/n, as expected. The increase in variability for the "All" estimator is smaller and bounded by the relatively good behaviour of the estimator when there are only individuals with missing y_3. The same pattern is found, in this case for both bias and variability, when $\alpha = 0.8$.

Table 2. Monte Carlo Estimation Results, $T = 5$, $n = 1000$

	n_3/n	Moments	Bias	Rmse	Med Bias	IQR
$\alpha = 0.4$	0		0.0028	0.0308	0.0016	0.0408
	0.2	CC	0.0029	0.0337	0.0027	0.0454
		All	0.0031	0.0316	0.0016	0.0430
	0.5	CC	0.0021	0.0444	0.0011	0.0564
		All	0.0018	0.0333	0.0013	0.0448
	0.8	CC	0.0092	0.0863	0.0054	0.0990
		All	0.0007	0.0357	0.0001	0.0445
	1		0.0033	0.0357	0.0013	0.0480
$\alpha = 0.8$	0		0.0440	0.1170	0.0131	0.1239
	0.2	CC	0.0452	0.1254	0.0151	0.1438
		All	0.0436	0.1231	0.0096	0.1322
	0.5	CC	0.0664	0.1520	0.0371	0.1992
		All	0.0487	0.1246	0.0179	0.1450
	0.8	CC	0.0863	0.1976	0.0601	0.2730
		All	0.0517	0.1305	0.0174	0.1731
	1		0.0488	0.1248	0.0194	0.1513

5 AR1X Model

Consider next the AR(1) model with an explanatory variable

$$y_{it} = \alpha y_{i,t-1} + \beta x_{it} + \eta_i + v_{it}.$$

Observations on the outcome y_{it} may be missing, but x_{it} is always observed.

As now $y_{it} = \alpha^p y_{i,t-p} + \beta \sum_{q=0}^{p-1} \alpha^q x_{i,t-q} + \eta_i \sum_{r=0}^{p-1} \alpha^r + \sum_{l=0}^{p-1} \alpha^l v_{i,t-l}$, and hence $y_{ij}^o = \alpha^{d_j} y_{i,j-1}^o + \beta \sum_{q=0}^{d_j-1} \alpha^q x_{i,t(j)-q} + \eta_i \sum_{r=0}^{d_j-1} \alpha^r + \sum_{l=0}^{d_j-1} \alpha^l v_{i,t(j)-l}$ we get

$$A(\alpha) y_i^o - \widetilde{x}_i \beta = \tau(\alpha) \eta_i + \widetilde{v}_i(\alpha),$$

where the $(j-1)$-th element of $\widetilde{x}_i$ is given by

$$\widetilde{x}_{i,j-1} = \sum_{q=0}^{d_j-1} \alpha^q x_{i,t(j)-q},$$

for $j = 2..., T^o$. The $(j-2)$-th element of the transformation $G(\alpha)\widetilde{x}_i$ is then given by

$$\sum_{q=0}^{d_j-1} \alpha^q x_{i,t(j)-q} - \frac{1-\alpha^{d_j}}{1-\alpha^{d_{j-1}}} \sum_{q=0}^{d_{j-1}-1} \alpha^q x_{i,t(j-1)-q},$$

for $j = 3, ..., T^o$, and hence the moment conditions (9) become

$$E\left[z_{i,,j-2}\left(\Delta y_{ij}^o - \alpha^{d_{j-1}} \frac{1-\alpha^{d_j}}{1-\alpha^{d_{j-1}}} \Delta y_{i,j-1}^o - \beta\left(\sum_{q=0}^{d_j-1} \alpha^q x_{i,t(j)-q} - \frac{1-\alpha^{d_j}}{1-\alpha^{d_{j-1}}} \sum_{q=0}^{d_{j-1}-1} \alpha^q x_{i,t(j-1)-q}\right)\right)\right] = 0,$$

for $j = 3, ..., T^o$, where $z_{i,j-2} = \left(y_{i1}^o, ..., y_{i,j-2}^o, x_{i1}, ..., x_{i,t(j)-e}\right)'$. e is determined by the endogeneity properties of x_{it}, e.g. $e = 1$ if x_{it} is predetermined: $E[x_{it}v_{it}] = E[x_{i,t-1}v_{i,t-1}] = 0$, but $E[x_{it}v_{i,t-1}] \neq 0$; and $e = 2$ if x_{it} is endogenous: $E[x_{it}v_{it}] \neq 0$. The Ahn-Schmidt equivalent moment conditions (10) follow straightforwardly.

References

[1] Abrevaya, J. (2013), The Projection Approach for Unbalanced Panel Data Models, *The Econometrics Journal* 16, 161-178.

[2] Ahn, S. and P. Schmidt (1995), Efficient Estimation of Models for Dynamic Panel Data, *Journal of Econometrics* 68, 5-27.

[3] Ahn, S. and P. Schmidt (1997), Efficient Estimation of Dynamic Panel Data Models: Alternative Assumptions and Simplified Estimation, *Journal of Econometrics* 76, 309-21.

[4] Albarran, P. and M. Arellano (2013), GMM Estimation from Incomplete and Rotating Panels, mimeo, CEMFI.

[5] Arellano, M. and S. Bond (1991), Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations, *Review of Economic Studies* 58, 277-97.

[6] Arellano, M. and O. Bover (1995). Another Look at the Instrumental Variable Estimation of Error-Components Models, *Journal of Econometrics 68*, 29-51.

[7] Arellano, M., O. Bover and J.M. Labeaga (1997), Autoregressive Models with Sample Selectivity for Panel Data, CEMFI Working Paper No. 9618.

[8] Blundell, R. and S. Bond (1998), Initial Conditions and Moment Restrictions in Dynamic Panel Data Models, *Journal of Econometrics 87*, 115-143.

[9] Kniesner, T., K. Viscusi, C. Woock and J. Ziliak (2012), The Value of Statistical Life: Evidence from Panel Data, *The Review of Economic and Statistics* 94, 74-87.

[10] Muris, C. (2013), Efficient GMM Estimation with a General Missing Data Pattern, mimeo, Simon Fraser University.

[11] Newey, W.K. and F. Windmeijer (2009), Generalized Method of Moments with Many Weak Moment Conditions, *Econometrica* 77, 687-719.

[12] Semykina, A. and J. Wooldridge (2013), Estimation of Dynamic Panel Data Models with Sample Selection, *Journal of Applied Econometrics* 28, 47-61.