UNIVERSITÄT PASSAU

Wirtschaftswissenschaftliche Fakultät

Practical aspects of using quadratic moment conditions in linear dynamic panel data models

Andrew Adrian Yu Pua, Markus Fritsch, Joachim Schnurbus

Diskussionsbeitrag Nr. B-38-19

Betriebswirtschaftliche Reihe ISSN 1435-3539

PASSAUER DISKUSSIONSPAPIERE

Herausgeber:
Die Gruppe der betriebswirtschaftlichen Professoren
der Wirtschaftswissenschaftlichen Fakultät
der Universität Passau
94030 Passau

Practical aspects of using quadratic moment conditions in linear dynamic panel data models

Andrew Adrian Yu Pua, Markus Fritsch, Joachim Schnurbus

Diskussionsbeitrag Nr. B-38-19

Betriebswirtschaftliche Reihe ISSN 1435-3539

Adresse des Autors/der Autoren:

Markus Fritsch
Wirtschaftswissenschaftliche Fakultät
Universität Passau
94030 Passau

Telefon: +49 851 509 2565
Telefax: +49 851 509 2562
E-Mail: markus.fritsch@uni-passau.de

Practical aspects of using quadratic moment conditions in linear dynamic panel data models

Andrew Adrian Yu Pua[1], Markus Fritsch[2], Joachim Schnurbus[3]

September 20, 2019

Abstract. We study the estimation of the lag parameter of linear dynamic panel data models with first order dynamics based on the quadratic Ahn and Schmidt (1995) moment conditions. Our contribution is twofold: First, we show that extending the standard assumptions by mean stationarity and time series homoscedasticity and employing these assumptions in estimation restores standard asymptotics and mitigates the non-standard distributions found in the literature. Second, we consider an IV estimator based on the quadratic moment conditions that consistently identifies the true population parameter under standard assumptions. Standard asymptotics hold for the estimator when the cross section dimension is large and the time series dimension is finite. We also suggest a data-driven approach to obtain standard errors and confidence intervals that preserves the time series dependence structure in the data.

Keywords. Panel data, linear dynamic model, quadratic moment conditions, root selection, standard asymptotics, inference.

JEL codes. C18, C23, C26.

[1]Wang Yanan Institute for Studies in Economics, Xiamen University, email: andrewypua@gmail.com
[2]University of Passau, Department of Statistics, email: markus.fritsch@uni-passau.de
[3]Chair of Statistics, University of Passau, email: joachim.schnurbus@uni-passau.de

1 Introduction

In this paper, we demonstrate the practical relevance of using the quadratic (in parameters) Ahn and Schmidt (1995) moment conditions in estimation of and inference on linear dynamic panel data models with first order dynamics for the case where the cross section dimension is large and the time series dimension is finite. We provide a practical perspective as to which regions of the parameter space and which elements of the data generating process (DGP) these moment conditions are likely to be more useful – especially when considered in isolation from other moment conditions.

Estimation of linear dynamic panel data models by the generalized method of moments (GMM) using a set of default linear (in parameters) moment conditions, such as those proposed by Anderson and Hsiao (1981), Holtz-Eakin, Newey, and Rosen (1988) and Arellano and Bover (1995), is popular in econometric practice due to the availability of routines in standard software (e.g., Roodman, 2009a; Kripfganz, 2018 for `Stata` and Croissant and Millo, 2008; Fritsch, Pua, and Schnurbus, 2019 for `R`). Ahn and Schmidt (1995) propose quadratic moment conditions that also arise from standard assumptions.[4] Incorporating these moment conditions in GMM estimation is not popular in practice because it requires nonlinear optimization. Unfortunately, Kiviet, Pleus, and Poldermans (2017) document extensively that the finite sample performance of the more popular linear GMM estimators is poor. Monte Carlo evidence from Bun and Sarafidis (2015) indicates that once the quadratic moment conditions are accounted for, finite sample performance may improve substantially.

Bun and Kleibergen (2014) and Gorgens, Han, and Xue (2016b) consider linear dynamic panel data models with first order dynamics and highlight the potential of using quadratic moment conditions to deal with identification failures that arise when employing the usual sets of linear moment conditions. In particular, Bun and Kleibergen (2014) modify the quadratic moment conditions to deal with the case where the true value of the lag parameter is unity. By considering worst case DGPs, they also show that GMM diagnostic testing may be affected and propose corrected versions of these tests. Gorgens, Han, and Xue (2016b) show that the quadratic moment conditions can provide full or partial identification of the lag parameter when the linear moment conditions fail to do so. They focus on characterizing the conditions for GMM identification (and failures thereof). Gorgens, Han, and Xue (2016a) derive the asymptotic distribution for a GMM estimator which employs linear and quadratic moment conditions under a set of standard

[4]Ahn and Schmidt (1997) propose a linearized GMM estimator that is asymptotically as efficient as the GMM estimator employing the quadratic moment conditions.

assumptions from the literature and the condition that point identification holds. They obtain non-standard distributions for DGPs besides the unit root case. We contribute to the literature as follows: First, we show that extending the standard assumptions by mean stationarity and time series homoscedasticity leads to point identification and retains standard asymptotics – even in the unit root case. Second, we consider the IV estimator of Pua, Fritsch, and Schnurbus (2019) based on quadratic moment conditions that consistently identifies the true population parameter under standard assumptions. The estimator incorporates a selection rule to choose from a set of candidate roots (or solutions) that does not depend on knowledge about the true population parameter or any population moment; standard asymptotics hold. Additionally, we propose a data-driven approach to compute standard errors and confidence intervals that preserves the time dependence structure of the data.

This paper is structured as follows: In Section 2, we introduce the modeling framework, discuss the practical relevance of existing results on quadratic moment conditions that point towards non-standard inference, and show how to restore standard asymptotics based on mean stationarity and time series homoscedasticity. We then sketch the IV estimator, the corresponding root selection rule, and the computation of standard errors and confidence intervals. Section 3 concludes.

2 Main results

2.1 Modeling framework

Consider the two equation form of the linear dynamic panel data model,

$$y_{i,t} = \rho_0 y_{i,t-1} + u_{i,t}, \qquad i = 1, \ldots, n;\; t = 2, \ldots, T, \tag{1}$$

$$u_{i,t} = c_i + \varepsilon_{i,t}, \tag{2}$$

where $y_{i,t}$ and $y_{i,t-1}$ are the dependent variable and its lag, ρ_0 is the true lag parameter, $u_{i,t}$ is a composite error term, c_i is an unobservable individual-specific effect, $\varepsilon_{i,t}$ is an idiosyncratic remainder component, i indicates the individual, t denotes the time dimension, and $t = 1$ is the initial time period.

We impose the assumptions:

(*A.1*) $\mathrm{E}[\varepsilon_{i,t} \cdot y_{i,1}] = 0, \qquad \forall i, t,$

(*A.2*) $\mathrm{E}[\varepsilon_{i,t} \cdot c_i] = 0, \qquad \forall i, t,$

(*A.3*) $\mathrm{E}[\varepsilon_{i,t} \cdot \varepsilon_{i,s}] = 0, \qquad \forall t \neq s,$

(*A.4*) $\rho_0 \in [-1, 1] \setminus \{0\},$

(*A.5*) $n \to \infty$, while T is fixed, such that $n/T \to 0$,

(*A.6*) $y_{i,1}$ is observable.

The Assumptions (*A.1*)-(*A.3*) require orthogonality of the idiosyncratic remainder components with the following model components: The initial conditions of the $y_{i,t}$-process, the unobservable individual-specific effects, and the idiosyncratic remainder components of other time periods. The latter requirement is also referred to as absence of serial correlation in the $\varepsilon_{i,t}$. Assumption (*A.4*) requires that $|\rho_0| \leq 1$ and that ρ_0 is non-zero. Assumption (*A.5*) requires that the cross section dimension is large, while the time series dimension is fixed, and Assumption (*A.6*) requires that the initial conditions are observable. Our assumptions are similar to the ones imposed in Ahn and Schmidt (1995); the only difference is that we also include the unit root case.

Plugging Equation (2) into Equation (1) yields the single equation form of the linear dynamic panel data model

$$y_{i,t} = \rho_0 y_{i,t-1} + c_i + \varepsilon_{i,t}, \qquad i = 1, \ldots, n; \; t = 2, \ldots, T. \tag{3}$$

First differencing the equation eliminates the unobserved individual-specific effect c_i; we obtain

$$\Delta y_{i,t} = \rho_0 \Delta y_{i,t-1} + \Delta u_{i,t}, \qquad \text{with} \qquad \Delta u_{i,t} = \Delta\varepsilon_{i,t}, \tag{4}$$
$$\text{and} \qquad \Delta\varepsilon_{i,t} = \varepsilon_{i,t} - \varepsilon_{i,t-1}.$$

Under the Assumptions (*A.1*)-(*A.6*), the linear moment conditions of Holtz-Eakin, Newey, and Rosen (1988) and the quadratic moment conditions of Ahn and Schmidt (1995) can be employed to estimate the lag parameter. In the following, we restrict our attention to estimation based on the quadratic moment conditions[5]

$$\mathrm{E}\left[u_{i,T} \cdot \Delta u_{i,t-1}\right] = 0, \qquad t = 4, \ldots, T. \tag{5}$$

[5]Compared to the originally proposed version of the moment conditions, we slightly adjust the time index of the first differenced composite error term for notational convenience.

Ahn and Schmidt (1995) note that the instruments $u_{i,T}$ are unobservable but can be represented in terms of parameters and observable model components. Rewriting the moment conditions in Equation (5), and using Assumption $(A.2)$ to drop the c_i when replacing the $u_{i,T}$, gives

$$\mathrm{E}\left[(y_{i,T} - \rho y_{i,T-1}) \cdot (\Delta y_{i,t-1} - \rho \Delta y_{i,t-2})\right] = 0. \tag{6}$$

The usual estimation strategy combines these quadratic moment conditions with linear moment conditions, like those proposed by Holtz-Eakin, Newey, and Rosen (1988), to form a GMM estimator and then employs nonlinear optimization techniques to obtain estimates of the lag parameter.

Instead of the GMM approach, we focus on the individual moment conditions in the first part of the paper. In the second part, we consider a simple IV estimator based on Equation (6), where the instrument is of the form $y_{i,T} - \rho y_{i,T-1}$. We obtain estimates of the lag parameter from solving the quadratic equation

$$\frac{1}{n}\sum_{i=1}^{n}\sum_{t=4}^{T}(y_{i,T} - \rho y_{i,T-1}) \cdot (\Delta y_{i,t-1} - \rho \Delta y_{i,t-2}) = 0. \tag{7}$$

2.2 Mean stationarity, homoscedasticity, and standard asymptotics

Theoretical analysis frequently imposes two auxiliary assumptions to simplify the derivations (see, e.g., Alvarez and Arellano, 2003; Hayakawa, 2015):

$(A.7)$ $\mathrm{E}[\Delta y_{i,t} c_i] = 0, \qquad \forall i, t,$

$(A.8)$ $\mathrm{E}[\varepsilon_{i,t}^2] = \sigma^2, \qquad \forall t,.$

The former assumption, which is commonly referred to as mean stationarity (see, e.g., Arellano, 2003), requires that the dependent variable and the unobserved individual-specific effects are constantly correlated over time. The latter assumption of time series homoscedasticity requires that the variance of the idiosyncratic remainder components is constant over time.

The Assumptions $(A.7)$ and $(A.8)$ provide additional restrictions that can be employed when estimating the lag parameter in Equation (3) based on the quadratic moment conditions. We consider the case where $T = 4$ to simplify the notation and assume that the Assumptions $(A.1)$-$(A.8)$ hold. From Equation (6), one moment condition is available for estimation. Rewriting the condition gives

$$\mathrm{E}\left[m_1(\rho)\right] = \mathrm{E}\left[y_{i,4}\Delta y_{i,3} - \rho(y_{i,3}\Delta y_{i,3} + y_{i,4}\Delta y_{i,2}) + \rho^2(y_{i,3}\Delta y_{i,2})\right] = 0. \tag{8}$$

The function $m_1(\cdot)$ is a quadratic polynomial in ρ.

Under the Assumptions $(A.1)$-$(A.6)$, the expected values of the terms corresponding to the three coefficients $1, \rho$, and ρ^2 of the quadratic polynomial can be represented by:

$$\begin{aligned}
\mathrm{E}\left[y_{i,4}\Delta y_{i,3}\right] =& \rho_0 \mathrm{E}\left[\varepsilon_{i,3}^2\right] + \rho_0^2(\rho_0-1)\mathrm{E}\left[\varepsilon_{i,2}^2\right] + \rho_0 \mathrm{E}\left[(c_i-(1-\rho_0)y_{i,1})y_{i,1}\right] \\
& + \rho_0\left(\rho_0^2+\rho_0+1\right)\mathrm{E}\left[(c_i-(1-\rho_0)y_{i,1})^2\right], \\
\mathrm{E}\left[y_{i,3}\Delta y_{i,3} + y_{i,4}\Delta y_{i,2}\right] =& \left(\rho_0^2+\rho_0(\rho_0-1)\right)\mathrm{E}\left[\varepsilon_{i,2}^2\right] + \mathrm{E}\left[\varepsilon_{i,3}^2\right] \\
& + (\rho_0+1)\mathrm{E}\left[(c_i-(1-\rho_0)y_{i,1})y_{i,1}\right] \\
& + \left(\rho_0(\rho_0+1)+(\rho_0^2+\rho_0+1)\right)\mathrm{E}\left[(c_i-(1-\rho_0)y_{i,1})^2\right], \\
\mathrm{E}\left[y_{i,3}\Delta y_{i,2}\right] =& \rho_0 \mathrm{E}\left[\varepsilon_{i,2}^2\right] + \mathrm{E}\left[(c_i-(1-\rho_0)y_{i,1})y_{i,1}\right] \\
& + (1+\rho_0)\mathrm{E}\left[(c_i-(1-\rho_0)y_{i,1})^2\right].
\end{aligned}$$

A helpful implication of Assumption $(A.7)$ for simplifying the three terms can be derived from the initial conditions restriction for a mean stationary process (see, e.g., Blundell and Bond, 1998):

$$y_{i,1} = \frac{c_i}{1-\rho_0} + v_{i,t}, \qquad \text{with} \qquad \mathrm{E}[v_{i,t}] = 0.$$

By rewriting the restriction, squaring the result, and taking expectations, it can be shown that the following expression holds:

$$\mathrm{E}\left[(c_i-(1-\rho_0)y_{i,1})^2\right] = 0.$$

It can then be derived from the Cauchy Schwarz inequality that

$$\left(\mathrm{E}\left[(c_i-(1-\rho_0)y_{i,1})y_{i,1}\right]\right)^2 \leq \mathrm{E}\left[(c_i-(1-\rho_0)y_{i,1})^2\right]\mathrm{E}\left[y_{i,1}^2\right] = 0. \tag{9}$$

Using this result and Assumption $(A.8)$, it can be shown that the three terms corresponding to the coefficients $1, \rho$, and ρ^2 of Equation (8) reduce to

$$\begin{aligned}
\mathrm{E}\left[y_{i,4}\Delta y_{i,3}\right] &= \rho_0\sigma^2 + \rho_0^2(\rho_0-1)\sigma^2, \\
\mathrm{E}\left[y_{i,3}\Delta y_{i,3} + y_{i,4}\Delta y_{i,2}\right] &= \left(\rho_0^2+\rho_0(\rho_0-1)\right)\sigma^2 + \sigma^2, \\
\mathrm{E}\left[y_{i,3}\Delta y_{i,2}\right] &= \rho_0\sigma^2
\end{aligned}$$

under mean stationarity and time series homoscedasticity.

We obtain the two roots (or solutions) from applying the quadratic formula to solve Equation (8) for ρ when imposing the Assumptions $(A.1)$-$(A.8)$:

$$\rho_{1/2} = \frac{\rho_0^2+\rho_0(\rho_0-1)+1 \pm |\rho_0-1|}{2\rho_0}, \qquad \text{with} \tag{10}$$

$$\rho_1 = \rho_0 \qquad \text{and} \qquad \rho_2 = \rho_0 - 1 + \frac{1}{\rho_0}.$$

The solution ρ_1 corresponds to the true population parameter. Note that since $\rho_0 - 1 \leq 0$ for $\rho_0 \in [-1, 1] \setminus \{0\}$, we are able to identify the true lag parameter – provided that we choose the solution which is closer to zero (i.e., the negative branch).

Although we achieve point identification of the estimator via a root selection rule that works asymptotically, standard asymptotics fail. To see the reason for this, consider the function underlying the quadratic moment conditions for $T = 4$:

$$m_1(\rho) = (y_{i,4} - \rho y_{i,3})(\Delta y_{i,3} - \rho \Delta y_{i,2}). \tag{11}$$

The first derivative of the function with respect to ρ is given by

$$\frac{\partial m_1(\rho)}{\partial \rho} = -(y_{i,3}\Delta y_{i,3} + y_{i,4}\Delta y_{i,2}) + 2\rho y_{i,3}\Delta y_{i,2}. \tag{12}$$

When the Assumptions (*A.1*)-(*A.8*) hold, computing the expected value of the first derivative yields

$$\sigma^2(-2\rho_0^2 + \rho_0 - 1 + 2\rho\rho_0). \tag{13}$$

In case $\rho = \rho_0$, this derivative simplifies to $-\sigma^2(1 - \rho_0)$ and equals zero in the unit root case (i.e., when $\rho_0 = 1$).

This violates one of the assumptions required to establish the standard results for the asymptotic distribution of the GMM estimator: The matrix containing the expected first derivatives of the functions underlying the moment conditions with respect to the model parameters needs to have full column rank (see, e.g., Proposition 7.10 in Hayashi, 2000 and Theorems 3.4 and 4.5 of Newey and McFadden, 1994) or, stated differently, the rank of the matrix needs to equal the length of the vector of estimated parameters. Note that as we consider a setting where we estimate only one parameter from a single moment condition in this section, the matrix of expected first derivatives reduces to a scalar, and this scalar is required to be non-zero to fulfill the rank condition. As mentioned after Equation (13), this condition does not hold in the unit root case under the Assumptions (*A.1*)-(*A.8*) and standard GMM asymptotics do not apply when estimating the lag parameter of the linear dynamic panel data model in Equation (3) based on the quadratic Ahn and Schmidt (1995) moment conditions.

2.3 Restoring standard asymptotics

The invalidity of standard GMM asymptotics sketched in Section 2.2 can be mitigated by deriving additional moment conditions from Assumption (*A.8*) and employing the conditions in the estimation of the lag parameter of Equation (3). Ahn and Schmidt (1995) derive moment conditions from time series homoscedasticity that are of the form

$$\mathrm{E}\left[\bar{u}_i \Delta u_{i,t}\right] = 0, \qquad t = 3, \ldots, T, \qquad \text{with} \tag{14}$$
$$\bar{u}_i = \frac{1}{T-1} \sum_{s=2}^{T} u_{i,s}.$$

The variable $\bar{u}_i$ denotes the time series average of the composite error terms. Similar to Equation (5), the time index in Equation (14) is slightly adjusted.

Consider again the setting where $T = 4$ and the Assumptions (*A.1*)-(*A.8*) hold. One of the additional moment conditions from imposing time series homoscedasticity is

$$\mathrm{E}\left[m_2(\rho)\right] = \mathrm{E}\left[\bar{u}_i \Delta u_{i,4}\right] = 0, \tag{15}$$

where the function $m_2(\cdot)$ denotes a function of ρ.

The lag parameter can then be estimated based on the two moment conditions $\mathrm{E}\left[m_1(\rho)\right]$ and $\mathrm{E}\left[m_2(\rho)\right]$. Employing the Assumptions (*A.1*)-(*A.8*) to simplify the conditions and computing the expectations yields the two restrictions

$$\rho_0\sigma^2 + \rho_0^2(\rho_0 - 1)\sigma^2 - \left(\left(\rho_0^2 + \rho_0(\rho_0 - 1)\right)\sigma^2 + \sigma^2\right)\rho + \rho^2\rho_0\sigma^2 = 0,$$
$$-(1-\rho)\sigma^2 - \rho_0\sigma^2 + \sigma^2 = 0.$$

The first restriction arises from $\mathrm{E}\left[m_1(\rho)\right]$ and the second from $\mathrm{E}\left[m_2(\rho)\right]$. Further simplifying the two equations and dividing both sides of both equations by σ^2 gives:

$$\rho_0 + \rho_0^2(\rho_0 - 1) - \left(\rho_0^2 + \rho_0(\rho_0 - 1) + 1\right)\rho + \rho^2\rho_0 = 0,$$
$$\rho - \rho_0 = 0.$$

The second equation shows that the lag parameter is point identified (i.e., $\rho = \rho_0$) by the restriction resulting from time series homoscedasticity.

As noted in Section 2.2, the matrix containing the expected first derivative of the functions underlying the moment conditions with respect to ρ needs to have full column rank. Since two moment conditions are available, and we only estimate the lag parameter here, the matrix reduces to a vector and can be represented by

$$\boldsymbol{G}(\rho) = \mathrm{E}\begin{pmatrix} \dfrac{\partial m_1(\rho)}{\partial \rho} \\ \dfrac{\partial m_2(\rho)}{\partial \rho} \end{pmatrix} = \begin{pmatrix} -\sigma^2(1-\rho_0) \\ \sigma^2 \end{pmatrix}. \tag{16}$$

The expected first derivatives are collected in $\boldsymbol{G}(\rho)$. Considering the two elements of the vector in greater detail reveals that the full column rank condition is now fulfilled for $\boldsymbol{G}(\rho)$ – as long as the variance of the idiosyncratic remainder components σ^2 is not equal to zero – and standard GMM asymptotics hold.

The practical implication of this result is that mean stationarity and time series homoscedasticity, and the restrictions implied by the assumptions, should be employed in the estimation of linear dynamic panel data models based on quadratic moment conditions, when the assumptions seem reasonable or when they are imposed by the study design anyway. This may be the case when building on theoretical work which uses these assumptions or when economic reasoning or theory indicate that the assumptions might be plausible. Given that the variance of the idiosyncratic remainder component is not equal to zero, using the assumptions assures point identification in the unit root case and ensures that standard asymptotics are valid. Note that these results are derived for first order dynamics and in the absence of non-lagged dependent explanatory variables. Generalizing the results may lead to further valuable insights.

2.4 Root selection in practice

Given the results in Section 2.3, it is natural to wonder if standard asymptotics can be established without imposing either of the two Assumptions (*A.7*) and (*A.8*) when estimating the lag parameter of Equation (3) based on the quadratic moment conditions. To investigate the matter, consider again the setting where $T = 4$ and impose only the Assumptions (*A.1*)-(*A.6*). Further define

$$M_1 = \mathrm{E}\left[(c_i - (1-\rho_0)y_{i,1})^2\right]$$

and

$$M_2 = \mathrm{E}\left[(c_i - (1-\rho_0)y_{i,1})y_{i,1}\right].$$

From solving the quadratic moment condition in Equation (6) for ρ with the quadratic formula, we obtain

$$\begin{aligned}\rho_{1/2} &= \frac{\mathrm{E}\left[y_{i,3}\Delta y_{i,3} + y_{i,4}\Delta y_{i,2}\right] \pm \sqrt{\left(\mathrm{E}\left[y_{i,3}\Delta y_{i,3} + y_{i,4}\Delta y_{i,2}\right]\right)^2 - 4\,\mathrm{E}\left[y_{i,4}\Delta y_{i,3}\right]\mathrm{E}\left[y_{i,3}\Delta y_{i,2}\right]}}{2\,\mathrm{E}\left[y_{i,3}\Delta y_{i,2}\right]} \\ &= \frac{(2\rho_0^2 - \rho_0)\sigma_2^2 + \sigma_3^2 + \left(2\rho_0^2 + 2\rho_0 + 1\right)M_1 + (\rho_0 + 1)M_2}{2\left(\rho_0\sigma_2^2 + (1+\rho_0)M_1 + M_2\right)} \\ &\quad \pm \frac{\left|\sigma_3^2 - \rho_0\sigma_2^2 + M_2(1-\rho_0) + M_1\right|}{2\left(\rho_0\sigma_2^2 + (1+\rho_0)M_1 + M_2\right)}.\end{aligned}$$

We interrupt solving the expression briefly at this step to note that the term

$$D = \sigma_3^2 - \rho_0\sigma_2^2 + M_2\left(1 - \rho_0\right) + M_1$$

has a neat interpretation. To see this, compute the expected first derivative of the function underlying the quadratic moment condition with respect to ρ and evaluate the derivative at $\rho = \rho_*$:

$$\mathrm{E}\left(\frac{\partial m_1(\rho)}{\partial \rho}\right)\Bigg|_{\rho=\rho_*} = -\left(\sigma_3^2 - \rho_*\sigma_2^2 + M_2(1-\rho_*) + M_1\right). \tag{17}$$

When evaluating the first derivative at the true population parameter (i.e., when setting $\rho_* = \rho_0$), the expression corresponds to $-D$.

Since the absolute value of D is taken when solving the quadratic equation with the quadratic formula, the roots of the quadratic moment condition obtained from applying the quadratic formula, regardless of the sign of D, are either $\rho_1 = \rho_0$ or

$$\begin{aligned}\rho_2 &= \frac{\sigma_3^2 + (\rho_0^2 - \rho_0)\sigma_2^2 + (\rho_0^2 + \rho_0 + 1)M_1 + M_2}{\rho_0\sigma_2^2 + (1+\rho_0)M_1 + M_2} \\ &= \rho_0 + \frac{\sigma_3^2 - \rho_0\sigma_2^2 + M_1 + M_2(1-\rho_0)}{\rho_0\sigma_2^2 + (1+\rho_0)M_1 + M_2} \\ &= \rho_0 + \frac{D}{\mathrm{E}\left[y_{i,3}\Delta y_{i,2}\right]}.\end{aligned}$$

Evaluating the expected first derivative in Equation (17) at this root (i.e., using $\rho_* = \rho_0 + D/\,\mathrm{E}[y_{i,3}\Delta y_{i,2}]$) coincides with $+D$.

Although the preceding results are neat, they are not sufficient to construct a root selection algorithm to choose between the two roots without any knowledge about the (sign of the) true population parameter or any population moments. In particular, when working with real world data sets, the expression inside the square root (i.e., D) could be negative – rendering the two roots to be complex conjugates of each other.

In order to establish a selection rule, first start by noting that when evaluating the first derivative in Equation (17) at $\rho_* = \rho_0$, we obtain

$$-\,\mathrm{E}\left[y_{i,3}\Delta y_{i,3} + y_{i,4}\Delta y_{i,2}\right] + 2\rho_0\,\mathrm{E}\left[y_{i,3}\Delta y_{i,2}\right],$$

while when evaluating at $\rho_* = D/\,\mathrm{E}[y_{i,3}\Delta y_{i,2}]$, we get

$$-\,\mathrm{E}\left[y_{i,3}\Delta y_{i,3} + y_{i,4}\Delta y_{i,2}\right] + 2\rho_0\,\mathrm{E}\left[y_{i,3}\Delta y_{i,2}\right] + 2D.$$

Then, consider subtracting the term $D/\,\mathrm{E}[y_{i,3}\Delta y_{i,2}]$ from the probability limit of both roots. We refer to these roots as adjusted roots. We either have

$$\rho_0 - \frac{D}{\mathrm{E}\left[y_{i,3}\Delta y_{i,2}\right]} = -\rho_0 + \frac{\mathrm{E}\left[y_{i,3}\Delta y_{i,3} + y_{i,4}\Delta y_{i,2}\right]}{\mathrm{E}[y_{i,3}\Delta y_{i,2}]}, \tag{18}$$

or

$$\rho_0 + \frac{D}{\mathrm{E}[y_{i,3}\Delta y_{i,2}]} - \frac{D}{\mathrm{E}[y_{i,3}\Delta y_{i,2}]} = \rho_0. \tag{19}$$

Therefore, we have established a root selection strategy: The root to be selected from the two candidate roots is the one where the root obtained from solving the quadratic formula matches the adjusted root. We propose the following algorithm:

1. Calculate the two roots of the quadratic equation. Refer to the negative branch of the roots of the quadratic formula as $\widehat{\rho}_1$ and to the positive branch as $\widehat{\rho}_2$. If the roots are complex conjugates of each other, take the absolute value of the discriminant and recalculate the roots.

2. Evaluate the sample counterpart of the expected first derivative of the function underlying the quadratic moment conditions with respect to ρ at the two roots obtained previously. Call them d_1 and d_2, respectively.

3. Compute $\widehat{\rho}_1 - d_1$ and $\widehat{\rho}_2 - d_2$. Choose $\widehat{\rho}_1$, if $\widehat{\rho}_2 - d_2 = \widehat{\rho}_1$; otherwise, choose $\widehat{\rho}_2$, if $\widehat{\rho}_1 - d_1 = \widehat{\rho}_2$.

Note that the proposed algorithm does not require any knowledge about the sign of D, the sign of $\mathrm{E}\left[y_{i,3}\Delta y_{i,2}\right]$, and ρ_0 to select one of the roots. As a result, we also do not need to determine which is the smaller or larger of the two roots for root selection purposes.

2.5 Implementation details

Computing the quadratic IV estimator given in Equation (7) is straightforward, as solving for the unknown parameter ρ in the quadratic equation is possible by using the quadratic formula. Given the advice in Section 2.4, directly computing the two roots only requires calculating certain sample averages from the data.

Since standard asymptotics apply as long as there is variation in the idiosyncratic remainder component, there are no separate limit laws when $\rho_0 < 1$ and $\rho_0 = 1$. We have the same root-n

rate of convergence for both cases. Furthermore, by not relying on the optimization of an objective function, no boundary issues arise since the asymptotic distribution of the consistent root may be directly obtained without reference to a local linear or Taylor series expansion. This allows for standard inference without any modifications.

We propose using the nonparametric bootstrap to compute standard errors and confidence intervals for the estimated parameter. Instead of sampling every observation with replacement, we sample the individual time series together for each cross sectional unit with replacement. This allows us to preserve the dependence structure in the time dimension of the panel. Bootstrap standard errors are obtained in the usual manner, while bootstrap $100\,(1-\alpha)\,\%$ confidence intervals are obtained as follows: First compute the $\alpha/2$ and $1-\alpha/2$ quantiles of the statistic formed by using the estimated root to center the bootstrapped roots. The lower and upper limit of the interval then result from subtracting the $1-\alpha/2$ and $\alpha/2$ quantiles from the estimated root.

3 Concluding remarks

We analyze the estimation of the lag parameter in linear dynamic panel data models with first order dynamics based on quadratic moment conditions. We employ standard assumptions from the literature and consider the setting where n is large and T is fixed and where the absolute value of the lag parameter is smaller or equal to one.

In the first part of the paper, we illustrate the reason for the non-standard distributions obtained in the literature for the unit root case: Standard GMM asymptotics are invalid, because the matrix containing the first derivatives of the functions underlying the quadratic moment conditions with respect to the parameter vector does not possess full column rank. We show that by imposing mean stationarity and time series homoscedasticity of the idiosyncratic remainder components and employing the implications of these assumptions in estimation, the lag parameter is point identified in the unit root case and standard GMM asymptotics can be retained. The requirement for this result to hold is that the variance of the idiosyncratic remainder components is not equal to zero.

In the second part of the paper, we consider an IV estimator based on quadratic moment conditions that does not require the assumptions of stationarity and time series homoscedasticity. In essence, parameter estimation boils down to solving a quadratic equation by the quadratic formula and yields two candidate roots (or solutions). The estimator incorporates a root selection rule that does not depend on knowledge about the true lag parameter or any population moment and chooses the correct root asymptotically. Thus, the true lag parameter is point identified by the estimator and standard asymptotics hold. Additionally, we propose to compute the standard errors and confidence intervals by a nonparametric bootstrap, where the individual time series (i.e., the individuals) are sampled instead of the individual observations, to preserve the dependence structure in the time dimension.

Our results shed light on the nature of the roots and which root is the consistent one. We believe that the alternative of looking into a GMM objective function is much less manageable. The results in this paper can be extended in different directions. Avenues for future research may be including non-lagged dependent explanatory variables into the analysis, considering higher-order dynamics, and/or generalizing the results to settings where n and T are large.

References

Ahn, SC and P Schmidt (1995). “Efficient estimation of models for dynamic panel data”. In: *Journal of Econometrics* 68.1, pp. 5–27.

Ahn, SC and P Schmidt (1997). “Efficient estimation of dynamic panel data models: Alternative assumptions and simplified estimation”. In: *Journal of Econometrics* 76.1–2, pp. 309–321.

Alvarez, J and M Arellano (2003). “The Time Series and Cross-Section Asymptotics of Dynamic Panel Data Estimators”. In: *Econometrica* 71.4, pp. 1121–1159.

Anderson, T and C Hsiao (1981). “Estimation of Dynamic Models with Error Components”. In: *Journal of the American Statistical Association* 76.375, pp. 598–606.

Arellano, M (2003). *Panel Data Econometrics*. Oxford University Press.

Arellano, M and O Bover (1995). “Another look at the instrumental variable estimation of error-components models”. In: *Journal of Econometrics* 68.1, pp. 29–51.

Blundell, R and S Bond (1998). “Initial conditions and moment restrictions in dynamic panel data models”. In: *Journal of Econometrics* 87.1, pp. 115–143.

Bun, MJG and F Kleibergen (2014). *Identification and inference in moments based analysis of linear dynamic panel data models*. UvA-Econometrics Discussion Paper 2013-07. Universiteit van Amsterdam, Dept. of Econometrics.

Bun, MJG and V Sarafidis (2015). “Chapter 3 – Dynamic Panel Data Models”. In: *The Oxford Handbook of Panel Data*. Ed. by BH Baltagi. Oxford University Press, pp. 76–110.

Croissant, Y and G Millo (2008). “Panel Data Econometrics in R: The plm Package”. In: *Journal of Statistical Software* 27.2, pp. 1–43.

Fritsch, M, A Pua, and J Schnurbus (2019). *pdynmc – An R-package for estimating linear dynamic panel data models based on nonlinear moment conditions*. University of Passau Working Papers in Business Administration B-39-19.

Gorgens, T, C Han, and S Xue (2016a). *Asymptotic distributions of the quadratic GMM estimator in linear dynamic panel data models*. ANU Working Papers in Economics and Econometrics 2016-635. Australian National University, College of Business and Economics, School of Economics.

Gorgens, T, C Han, and S Xue (2016b). *Moment restrictions and identification in linear dynamic panel data models*. ANU Working Papers in Economics and Econometrics 2016-633. Australian National University, College of Business and Economics, School of Economics.

Hayakawa, K (2015). "The Asymptotic Properties Of The System GMM Estimator In Dyanmic Panel Data Models When Both N And T Are Large". In: *Econometric Theory* 31.3, pp. 647–667.

Hayashi, F (2000). *Econometrics*. Princeton University Press.

Holtz-Eakin, D, K Newey Whitney, and HS Rosen (1988). "Estimating Vector Autoregressions with Panel Data". In: *Econometrica* 56.6, pp. 1371–1395.

Kiviet, JF, M Pleus, and R Poldermans (2017). "Accuracy and Efficiency of Various GMM Inference Techniques in Dynamic Micro Panel Data Models". In: *Econometrics* 5.1, p. 14.

Kripfganz, S (2018). *XTDPDGMM: Stata module to perform generalized method of moments estimation of linear dynamic panel data models*. Version 1.1.1. URL: http://EconPapers.repec.org/RePEc:boc:bocode:s458395.

Newey, WK and DL McFadden (1994). "Chapter 36 – Large sample estimation and hypothesis testing". In: ed. by RF Engle and DL MacFadden. Vol. 4. Handbook of Econometrics. Elsevier, pp. 2111–2245.

Pua, A, M Fritsch, and J Schnurbus (2019). *Large sample properties of an IV estimator based on the Ahn and Schmidt moment conditions*. University of Passau Working Papers in Business Administration B-37-19.

Roodman, D (2009a). "How to do xtabond2: An Introduction to Difference and System GMM in Stata". In: *Stata Journal* 9.1, pp. 86–136.

Betriebswirtschaftliche Reihe der Passauer Diskussionspapiere

Bisher sind erschienen:

B-1-98 Jochen Wilhelm, A fresh view on the Ho-Lee model of the term structure from a stochastic discounting perspective

B-2-98 Bernhard Nietert und Jochen Wilhelm, Arbitrage, Pseudowahrscheinlichkeiten und Martingale - ein didaktisch einfacher Zugang zur elementaren Bewertungstheorie für Derivate

B-3-98 Bernhard Nietert, Dynamische Portfolio-Selektion unter Berücksichtigung von Kurssprüngen

B-4-99 Jochen Wilhelm, Option Prices with Stochastic Interest Rates – Black/Scholes and Ho/Lee unified

B-5-99 Anja Tuschke, Stock-based Incentive Schemes and the Managerial Labor Market

B-6-00 Jochen Wilhelm, Das Gaußsche Zinsstrukturmodell – Eine Analyse auf der Basis von Wahrscheinlichkeitsverteilungen

B-7-01 Bernhard Nietert und Jochen Wilhelm, Some Economic Remarks on Arbitrage Theory

B-8-01 Jochen Wilhelm, Option Prices with Stochastic Interest Rates – Black/Scholes and Ho/Lee unified

B-9-02 Jochen Wilhelm, Risikoabschläge, Risikozuschläge und Risikoprämien – Finanzierungstheoretische Anmerkungen zu einem Grundproblem der Unternehmensbewertung

B-10-03 Jochen Wilhelm, Unternehmensbewertung – Eine finanzmarkttheoretische Untersuchung

B-11-04 Bernhard Nietert und Jochen Wilhelm, Non-Negativity of Nominal and Real Riskless Rates, Arbitrage Theory, and the Null-Alternative Cash

B-12-06 Armin Dolzer und Bernhard Nietert – Portfolio selection with Time Constraints and a Rational Explanation of Insufficient Diversification and Excessive Trading

B-13-08 Josef Schosser - Bewertung ohne "Kapitalkosten": ein arbitragetheoretischer An satz zu Unternehmenswert, Kapitalstruktur und persönlicher Besteuerung

B-14-14 Mathias Eickholt, Oliver Entrop, Marco Wilkens, Individual Investors and Suboptimal Early Exercises in the Fixed-Income Market

B-15-14 Mathias Eickholt, Oliver Entrop, Marco Wilkens, What Makes Individual Investors Exercise Early? Empirical Evidence from the Fixed-Income Market

B-16-14 Mathias Eickholt, Behavioral Financial Engineering in the Fixed-Income Market: The Influence of the Coupon Structure

B-17-16 Caroline Baethge, Performance in the Beauty Contest: How Strategic Discussion Enhances Team Reasoning

B-18-16 Caroline Baethge, Marina Fiedler, Aligning Mission Preferences: Does Self-Selection Foster Performance in Working Groups?

B-19-16 Caroline Baethge, Marina Fiedler, All or (Almost) Nothing? The Influence of Information Cost and Training on Information Selection and the Quality of Decision-Making.

B-20-16 Caroline Baethge, Marina Fiedler, Ernan Haruvey, In It to Win It: Experimental Evidence on Unique Bid Auctions

B-21-16 Markus Grottke, Maximilian Kittl, First the stick, then the carrot? A cross-country evaluation of the OECD's initiative against harmful tax competition

B-22-16 Heike Diller, Stephen Jeffrey, Marina Fiedler, Searching for the silver linings of techno-invasion

B-23-16 Stephen Jeffrey, Heike Diller, Marina Fiedler, How does intensification and mobile rearrangement affect employee commitment

B-24-16 Heike Diller, Life is tough so you gotta be rough – How resilience impacts employees' attitude towards ICT use

B-25-16 Stephen Jeffrey, Heike Diller, Marina Fiedler, Closing the Strategy-Performance Gap: The Role of Communication Fit and Distraction

B-26-17 S. Baller, O. Entrop, A. Schober, M. Wilkens, What drives Performance in the Speculative Market of Short-Term Exchange-Traded Retail Products?

B-27-17 S. Baller, Risk Taking in the Market of Speculative Exchange-Traded Retail Products: Do Socio-Economic Factors matter?

B-28-17 L. Frey, L. Engelhard, Review on Tax Research in Accounting: Is the information given by U.S. GAAP income taxes also provided by IFRS?

B-29-17 J. Lorenz, M. Diller, Do Tax Information Exchange Agreements Curb Transfer Pricing-Induced Tax Avoidance?

B-30-17 J. Lorenz, M. Grottke, Tax Consultants' Incentives – A Game-Theoretic Investigation into the Behavior of Tax Consultants, Taxpayers and the Tax Authority in a Setting of Tax Complexity

B-31-18 Oliver Entrop, Matthias F. Merkel, "Exchange Rate Risk" within the European Monetary Union? Analyzing the Exchange Rate Exposure of German Firms

B-32-18 Oliver Entrop, Matthias F. Merkel, Manager's Research Education, the Use of FX Derivatives and Corporate Speculation

B-33-18 Matthias F. Merkel, Foreign Exchange Derivative Use and Firm Value: Evidence from German Non-Financial Firms

B-34-19 Oliver Entrop, Georg Fischer, Hedging Costs and Joint Determinants of Premiums and Spreads in Structured Financial Products

B-35-19 Georg Fischer, How Dynamic Hedging Affects Stock Price Movements: Evidence from German Option and Certificate Markets

B-36-19 Markus Fritsch, On GMM estimation of linear dynamic panel data models

B-37-19 Andrew Adrian Yu Pua, Markus Fritsch, Joachim Schnurbus, Large sample properties of an IV estimator based on the Ahn and Schmidt moment conditions