Estimating dynamic panel data models: a guide for macroeconomists

Ruth A. Judson\textsuperscript{a}, Ann L. Owen\textsuperscript{b,}\textsuperscript{*}

\textsuperscript{a}Federal Reserve Board of Governors, 20th & C Sts., N.W. Washington, D.C. 20551, USA
\textsuperscript{b}Hamilton College, 198 College Hill Road, Clinton, NY 13323, USA

Received 11 December 1998; accepted 12 May 1999

Abstract

Using a Monte Carlo approach, we find that the bias of LSDV for dynamic panel data models can be sizeable, even when $T = 20$. A corrected LSDV estimator is the best choice overall, but practical considerations may limit its applicability. GMM is a second best solution and, for long panels, the computationally simpler Anderson–Hsiao estimator performs well. © 1999 Elsevier Science S.A. All rights reserved.

Keywords: Panel data; Simulation; Dynamic model

JEL classification: C23; O11; E00

1. Introduction

The revitalization of interest in long-run growth and the availability of macroeconomic data for large panels of countries has generated interest among macroeconomists in estimating dynamic panel models. However, microeconomists have generally been more avid users of panel data, and, thus, existing panel techniques have been devised and tested with the typical dimensions of microeconomic datasets in mind. These datasets usually have a time dimension far smaller and an individual (country) dimension far greater than the typical macroeconomic panel.

This difference is important in choosing an estimation technique for two reasons. First, it is well known that the LSDV (least squares dummy variable) model with a lagged dependent variable generates biased estimates when the time dimension of the panel ($T$) is small. Thus, for many

\textsuperscript{*}Corresponding author. Tel.: +1-315-859-4419; fax: +1-315-859-4477.
\textit{E-mail address:} aowen@hamilton.edu (A.L. Owen)

0165-1765/99/$ – see front matter © 1999 Elsevier Science S.A. All rights reserved.
PII: S0165-1765(99)00130-5
macroeconomists, the question, ‘How big should $T$ be before the bias can be ignored?’ is a critical one. A second reason that macro panels may require different estimation techniques than those used on micro panels is that recent work investigating the appropriateness of competing estimators has generated conflicting results, showing that the characteristics of the data influence the performance of an estimator.\footnote{Arellano and Bond (1991) find that GMM procedures are more efficient than the Anderson–Hsiao estimator. However, Kiviet (1995), using a slightly different experimental design, finds that the Anderson–Hsiao estimator compares favorably to GMM and concludes that no estimator is appropriate in all circumstances.}

We evaluate several different techniques for estimating dynamic models with panels characteristic of many macroeconomic panel datasets; our goal is to provide a guide to choosing appropriate techniques for panels of various dimensions. Our work most closely follows Kiviet’s (1995); however, we focus our attention on data with the qualities normally encountered by macroeconomists while he focuses on the short (small $T$), wide (large $N$) panels typical of micro data.\footnote{We confine ourselves to models and techniques most likely to be of practical use in macro panels. We thus limit our study to stationary data, GMM with first-moment instruments, and $T$ between 5 and 30. Other Monte Carlo studies have considered models beyond these parameters: Kao (1997) and Pedroni (1997) explore unit roots; Blundell and Bond (1998) investigate restricting initial conditions; Ziliak (1997) reviews a variety of methods, but only for $T \leq 11$. Ahn and Schmidt (1995) and Keane and Runkle (1992) consider exploiting additional moment restrictions. Since the writing of this paper, Bun and Kiviet (1998) have done further work in which they consider moderate $T$ and $N$.}

We have three main conclusions. First, macroeconomists should not dismiss the LSDV bias as insignificant. Even with a time dimension as large as 30, we find that the bias may be equal to as much as 20% of the true value of the coefficient of interest. However, using an RMSE criterion, the LSDV performs just as well or better than many alternatives when $T=30$. With a smaller time dimension, LSDV does not dominate the alternatives. Second, for panels of all sizes, a corrected LSDV estimator generally has the lowest RMSE. However, implementation of the corrected LSDV for an unbalanced panel has not been derived and therefore alternatives may be needed. When the corrected LSDV is not practical, a GMM procedure usually produces lower RMSEs relative to the Anderson–Hsiao estimator. Finally, we find that a ‘restricted GMM’ estimator that uses a subset of the available lagged values as instruments increases computational efficiency without significantly detracting from its effectiveness.

2. The problem and proposed solutions

We consider the dynamic fixed effects model

$$y_{i,t} = \gamma y_{i,t-1} + x_{i,t}' \beta + \eta_i + e_{i,t}; |\gamma| < 1$$

(1)

where $\eta_i$ is a fixed-effect, $x_{i,t}$ is a $(K-1) \times 1$ vector of exogenous regressors and $e_{i,t} \sim N(0, \sigma^2_e)$ is a random disturbance. We assume

$$\sigma^2_e > 0,$$

$$E(e_{i,t}, e_{j,s}) = 0 \quad i \neq j \text{ or } t \neq s$$

$$E(x_{i,t}, e_{i,s}) = 0 \quad \forall \ i, j, t, s$$

(2)
The fixed effects model we have chosen is a common choice for macroeconomists, and is generally more appropriate than a random effects model for two reasons. First, if the individual effect represents omitted variables, it is likely that these country-specific characteristics are correlated with the other regressors. Second, it is also likely that a typical macro panel will contain most countries of interest and, thus, will not likely be a random sample from a much larger universe of countries (e.g. an OECD panel contains most OECD countries).

The model in Eq. (1), however, includes as one of the regressors a lagged dependent variable. In this case, the usual approach to estimating a fixed-effects model, LSDV, generates a biased estimate of the coefficients. Nickell (1981) derives an expression for the bias of $\gamma$ when there are no exogenous regressors, showing that the bias approaches zero as $T$ approaches infinity. Thus, LSDV only performs well when the time dimension of the panel is large.

Several estimators have been proposed to estimate Eq. (1) when $T$ is moderate. With a typical macro dataset in mind, we implement a Monte Carlo study to consider four estimators: an instrumental variables estimator proposed by Anderson and Hsiao (1981), two GMM estimators discussed in Arellano and Bond (1991), and a corrected LSDV estimator derived in Kiviet, 1995. Henceforth, we call the Anderson–Hsiao estimator, AH, Arellano and Bond’s one-step estimator GMM1 and their two-step estimator GMM2, and Kiviet’s corrected LSDV estimator, LSDVC.

3. Methodology

Our data generation process closely follows Kiviet (1995). The model for $y_{it}$ is given in Eq. (1); $x_{it}$ was generated with

$$x_{it} = \rho x_{i,t-1} + \xi_{i,t} \sim N(0, \sigma_\xi^2).$$

(3)

Thus, in addition to $\beta$, $\rho$ and $\sigma_\xi^2$ also determine the correlation between $y_{it}$ and $x_{it}$. Kiviet defines a signal to noise ratio, $\sigma_s^2$,

$$\sigma_s^2 = \var(v_{it} - e_{it}), \quad v_{it} = y_{i,t} - \frac{1}{1 - \gamma} \eta_t$$

(4)

and shows that it can be calculated from other parameters of the model as follows

$$\sigma_s^2 = \beta^2 \sigma_e^2 \left[ 1 + \left( \frac{\gamma + \rho}{1 + \gamma \rho} \right)^2 \frac{1}{\gamma \rho - 1} - (\gamma \rho)^2 \right]^{-1} + \frac{\gamma^2}{1 - \gamma} \sigma_e^2.$$  

(5)

The higher the signal-to-noise ratio, the more useful $x_{it}$ is in explaining $y_{it}$. Kiviet (1995) finds that varying the signal-to-noise ratio significantly alters the relative bias of the estimators.

We also choose $\beta = 1 - \gamma$ so that a change in $\gamma$ affects only the short-run dynamic relationship

---

3 We consider here only the AH estimator that uses the lagged level as an instrument because Arellano (1989) shows that using the lagged difference is inefficient.

4 Anderson and Hsiao (1981), Arellano and Bond (1991), and Kiviet (1995) offer a more thorough discussion of each of these estimators. GAUSS programs are available from the authors.
between \( x \) and \( y \) and not the steady-state relationship. Thus, given choices for \( \gamma, \sigma_x^2, \sigma_y^2, \) and \( \rho, \) all of the other parameters of the model are determined. Our parameter choices can be summarized as follows: \( \sigma_x^2 \) is normalized to 1, \( \rho \) is set at the intermediate value of 0.5, \( \sigma_y^2 \) alternates between a value of 2 and 8, and \( \gamma \) alternates between 0.2 and 0.8. For each combination of parameters we vary the size of our panel. \( N, \) the cross-sectional dimension, takes on values of 20 or 100, and \( T, \) the time dimension, is assigned values of 5, 10, 20 and 30. In total, we have 32 different parameter combinations.

We generate the data by choosing \( x_{i,0}, y_{i,0} = 0 \) and then discarding the first 50 observations before selecting our sample. We performed 1000 replications with fixed seeds for the random number generator so that our results can be replicated.

4. Results

We first examine the bias of the OLS and LSDV estimators for various panel sizes. Table 1 summarizes the results from this initial experiment for a subset of parameter values.\(^5\) These results confirm several well-documented conclusions about these estimators: (1) in both cases, the bias of \( \gamma \) is more severe than that for \( \beta, \) (2) OLS provides biased estimates even for large \( T, \) and (3) the bias of the LSDV estimator increases with \( \gamma \) and decreases with \( T. \) In addition, these results show that the bias of the LSDV estimate is not unsubstantial, even at \( T=20. \) When \( T=30, \) the average bias becomes smaller, although the LSDV does not become more efficient. Based on the results in Table 1, one could expect an LSDV estimate with a bias from 3% to 20% of the true value of the coefficient even when \( T=30. \) Errors of this magnitude, however, would still result in an estimate with the correct sign.

Since LSDV is often inappropriate, we explore the properties of other estimators. Before making an overall comparison, we first narrow our selection by comparing various GMM procedures. Arellano...

---

Table 1
OLS and LSDV bias estimates\(^6\)

<table>
<thead>
<tr>
<th>( T )</th>
<th>( \gamma )</th>
<th>( \gamma ) bias</th>
<th>LSDV (S.E.)</th>
<th>( \beta )</th>
<th>( \beta ) bias</th>
<th>LSDV (S.E.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>0.2</td>
<td>0.225 (0.039)</td>
<td>-0.147 (0.040)</td>
<td>0.8</td>
<td>-0.098 (0.044)</td>
<td>0.006 (0.045)</td>
</tr>
<tr>
<td>10</td>
<td>0.2</td>
<td>0.049 (0.026)</td>
<td>-0.504 (0.058)</td>
<td>0.2</td>
<td>-0.005 (0.055)</td>
<td>-0.027 (0.070)</td>
</tr>
<tr>
<td>20</td>
<td>0.2</td>
<td>0.225 (0.032)</td>
<td>-0.059 (0.023)</td>
<td>0.8</td>
<td>-0.099 (0.031)</td>
<td>0.015 (0.026)</td>
</tr>
<tr>
<td>30</td>
<td>0.2</td>
<td>0.049 (0.017)</td>
<td>-0.232 (0.032)</td>
<td>0.2</td>
<td>-0.007 (0.037)</td>
<td>0.002 (0.045)</td>
</tr>
<tr>
<td></td>
<td>0.8</td>
<td>0.225 (0.028)</td>
<td>-0.027 (0.015)</td>
<td>0.2</td>
<td>-0.100 (0.023)</td>
<td>0.009 (0.017)</td>
</tr>
<tr>
<td></td>
<td>0.8</td>
<td>0.049 (0.012)</td>
<td>-0.104 (0.019)</td>
<td>0.2</td>
<td>-0.008 (0.026)</td>
<td>0.006 (0.028)</td>
</tr>
</tbody>
</table>

---

\(^5\) Results for a full set of parameter values for this experiment and all others described in this paper are qualitatively similar to those reported and are available from the authors.

---

\(^6\) 1000 draws; \( N = 100; \sigma_x = 1; \sigma_y = 2; \rho = 0.5. \)
and Bond (1991) discuss two variants of a GMM procedure that use all lagged values as instruments. When $T$ gets large, however, computational requirements increase substantially and a GMM estimation using all available instruments may not be practical to implement. Results from Monte Carlo simulations (available from the authors) indicate that (1) the one-step GMM estimator outperforms the two-step estimation, and (2) using a 'restricted GMM' procedure (the number of values of the lagged dependent variable and the exogenous regressors used as instruments is reduced to three, five, and seven instruments) does not materially reduce the performance of this technique.\(^6\)

Finally, we make an overall comparison between OLS, LSDV, AH, GMM and LSDVC. Based on the GMM comparisons reported above, we focus on two restricted GMM estimators — GMM13 and GMM17 which are GMM1 with three and seven instruments, respectively. Table 2 shows the average bias, standard deviations and RMSEs of our estimates of $\gamma$ (the bias of the estimates of $\beta$ are relatively small and cannot be used to distinguish between estimators).

The results in Table 2 show that all the estimators (with the exception of OLS) generally perform better with a larger $N$ and $T$. Thus, for a sufficiently large $N$ and $T$, the differences in efficiency, bias and RMSEs of the different techniques become quite small. Even so, the results in Table 2 do highlight one technique that consistently outperforms the others — LSDVC.

Unfortunately, while LSDVC may produce superior results, it is not always practical to implement. In particular, a method of implementing LSDVC for an unbalanced panel has not yet been developed. This is a particularly important consideration for macro data since the likelihood of having an unbalanced panel may increase as the time dimension gets large. Our results indicate that if LSDVC cannot be implemented that (1) when $T=30$, LSDV performs just as well or better than the viable alternatives, (2) when $T \leq 10$, GMM is best and (3) when $T=20$, GMM or AH may be chosen. While GMM still produces the lowest RMSEs even when $T=20$, the difference in performance is not that great and computational issues may become more important. Because the efficiency of the AH estimator increases substantially as $T$ gets larger, the computationally simpler AH method may be justified when $T$ is large enough.

### 5. Conclusion

The recommendations from our Monte Carlo analysis are summarized below.

<table>
<thead>
<tr>
<th>Summary of recommendations</th>
<th>$T \leq 10$</th>
<th>$T=20$</th>
<th>$T=30$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Balanced panel</td>
<td>LSDVC</td>
<td>LSDVC</td>
<td>LSDVC</td>
</tr>
<tr>
<td>Unbalanced panel</td>
<td>GMM1</td>
<td>GMM1 or AH</td>
<td>LSDV</td>
</tr>
</tbody>
</table>

\(^6\)For example, when we use seven instruments, we use seven lagged values of the dependent variable (if available). For the exogenous regressors, $x$, we use the seven closest values — three previous values, the contemporaneous value and three future values (when available). We also checked if there were gains to iterating GMM or to using more than seven instruments for $T=20$ (where 18 instruments would be available). There was minimal gain to either of these procedures.
Table 2
Bias estimates for $\gamma$ using various estimators$^a$

<table>
<thead>
<tr>
<th>$T$</th>
<th>$N$</th>
<th>$\gamma$</th>
<th>OLS (S.E.)</th>
<th>LSDV (S.E.)</th>
<th>LSDVC (S.E.)</th>
<th>A-HIV (S.E.)</th>
<th>GMM13 (S.E.)</th>
<th>GMM17 (S.E.)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>RMSE</td>
<td>RMSE</td>
<td>RMSE</td>
<td>RMSE</td>
<td>RMSE</td>
<td>RMSE</td>
</tr>
<tr>
<td>5</td>
<td>20</td>
<td>0.2</td>
<td>0.207 (0.148)</td>
<td>0.001 (0.017)</td>
<td>-0.052 (0.017)</td>
<td>-0.052 (0.017)</td>
<td>-0.069 (0.017)</td>
<td>-0.069 (0.017)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.8</td>
<td>0.033 (0.507)</td>
<td>-0.190 (0.061)</td>
<td>-0.357 (0.044)</td>
<td>-0.357 (0.044)</td>
<td>-0.404 (0.044)</td>
<td>-0.404 (0.044)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>100</td>
<td>0.225 (-0.147)</td>
<td>-0.006 (0.000)</td>
<td>-0.011 (0.015)</td>
<td>-0.011 (0.015)</td>
<td>-0.041 (0.015)</td>
<td>-0.041 (0.015)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.8</td>
<td>0.049 (-0.004)</td>
<td>-0.131 (0.007)</td>
<td>-0.116 (0.154)</td>
<td>-0.116 (0.154)</td>
<td>-0.154 (0.154)</td>
<td>-0.154 (0.154)</td>
</tr>
<tr>
<td>10</td>
<td>20</td>
<td>0.2</td>
<td>0.210 (-0.060)</td>
<td>0.000 (0.005)</td>
<td>-0.038 (0.048)</td>
<td>-0.038 (0.048)</td>
<td>-0.050 (0.048)</td>
<td>-0.050 (0.048)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.8</td>
<td>0.038 (-0.238)</td>
<td>-0.049 (0.013)</td>
<td>-0.214 (0.236)</td>
<td>-0.214 (0.236)</td>
<td>-0.236 (0.236)</td>
<td>-0.236 (0.236)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>100</td>
<td>0.225 (-0.059)</td>
<td>0.000 (0.000)</td>
<td>-0.009 (0.011)</td>
<td>-0.009 (0.011)</td>
<td>-0.011 (0.011)</td>
<td>-0.011 (0.011)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.8</td>
<td>0.049 (-0.232)</td>
<td>-0.032 (0.000)</td>
<td>-0.056 (0.081)</td>
<td>-0.056 (0.081)</td>
<td>-0.081 (0.081)</td>
<td>-0.081 (0.081)</td>
</tr>
<tr>
<td>20</td>
<td>20</td>
<td>0.2</td>
<td>0.213 (-0.028)</td>
<td>-0.001 (0.001)</td>
<td>-0.033 (0.036)</td>
<td>-0.033 (0.036)</td>
<td>-0.036 (0.036)</td>
<td>-0.036 (0.036)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.8</td>
<td>0.041 (-0.108)</td>
<td>-0.007 (0.003)</td>
<td>-0.130 (0.143)</td>
<td>-0.130 (0.143)</td>
<td>-0.143 (0.143)</td>
<td>-0.143 (0.143)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>100</td>
<td>0.225 (0.027)</td>
<td>0.000 (0.001)</td>
<td>-0.007 (0.008)</td>
<td>-0.007 (0.008)</td>
<td>-0.008 (0.008)</td>
<td>-0.008 (0.008)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.8</td>
<td>0.049 (-0.104)</td>
<td>-0.005 (0.001)</td>
<td>-0.029 (0.040)</td>
<td>-0.029 (0.040)</td>
<td>-0.040 (0.040)</td>
<td>-0.040 (0.040)</td>
</tr>
<tr>
<td>30</td>
<td>20</td>
<td>0.2</td>
<td>0.214 (-0.018)</td>
<td>0.000 (0.001)</td>
<td>-0.031 (0.032)</td>
<td>-0.031 (0.032)</td>
<td>-0.032 (0.032)</td>
<td>-0.032 (0.032)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.8</td>
<td>0.043 (-0.068)</td>
<td>-0.001 (0.003)</td>
<td>-0.108 (0.115)</td>
<td>-0.108 (0.115)</td>
<td>-0.115 (0.115)</td>
<td>-0.115 (0.115)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>100</td>
<td>0.226 (-0.017)</td>
<td>0.000 (0.000)</td>
<td>-0.007 (0.007)</td>
<td>-0.007 (0.007)</td>
<td>-0.007 (0.007)</td>
<td>-0.007 (0.007)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.8</td>
<td>0.049 (-0.066)</td>
<td>-0.001 (0.000)</td>
<td>-0.024 (0.030)</td>
<td>-0.024 (0.030)</td>
<td>-0.030 (0.030)</td>
<td>-0.030 (0.030)</td>
</tr>
</tbody>
</table>

$^a$ 1000 draws; $\sigma_x = 1; \sigma_\varepsilon = 2; \rho = 0.5; \mu = 1.$
References