Indirect Estimation of Panel Models With Time Varying Latent Components

Ivana Komunjer* Serena Ng†

August 2010

Comments Welcome

Abstract

This paper considers an indirect inference approach that exploits the biases in an auxiliary model to identify the parameters of interests. The proposed augmented indirect inference estimator (IDEA) is non-standard because (i) the covariates cannot be held fixed in simulations, and (ii) the auxiliary parameters must be chosen to vary with the nuisance parameters causing inconsistency. We provide a Panel-ME algorithm for mismeasured dynamic panel models that does not require fully specifying the joint distribution of the data. A simple modification leads to a Panel-IV algorithm for models with endogenous variables. A Panel-CS algorithm is also proposed for dynamic panel models with cross-section dependence which, like measurement error models, also have time varying latent components. A Monte Carlo study shows that all three algorithms have impressive finite sample properties.

JEL Classification: C1, C3

Keywords: Dynamic Panel, Fixed Effects, Measurement Error, Endogenous Regressors, Cross-Section Dependence.

*University of California, San Diego, 9500 Gilman Drive, La Jolla, CA 92093. Email: komunjer@ucsd.edu

†Columbia University, 420 W. 118 St. MC 3308, New York, NY 10027. Email: serena.ng@columbia.edu

This paper was presented at the 2010 International Panel Data Conference in Amsterdam. The authors acknowledge financial support from the National Science Foundation SES-0962473 and SES-0962431.

1 Introduction

Few results are available in the literature for dynamic panel models with time varying latent components of which serially correlated measurement errors is a special case. Instead of working around the measurement errors to look for orthogonality conditions, we take as a starting point that the biases induced by measurement errors contain valuable information about the parameters of interest. We use the estimates of an auxiliary model that are known to be inconsistent to obtain consistent estimates of the original model. Our approach falls in the class of Indirect Inference methods proposed by Gourieroux, Monfort, and Renault (1993) and Smith (1993). The typical indirect inference estimator (IDE) either holds the covariates fixed in simulations, or fully specifies their data generating processes so that they can be simulated along with the endogenous variables. We enhance the IDE to handle situations when the parameters in the marginal distribution of the covariates and those of the conditional model are not variation free. The proposed estimator will hereafter be referred to as IDEA (augmented indirect inference estimator).

Our IDEA estimator is a promising alternative to least squares estimation, which is convenient and often efficient when the regression error is orthogonal to the covariates. However, in the case of panel data, it is well-known that least squares estimation using demeaned or differenced data is inconsistent when the time dimension of the panel is short. Furthermore, if data are measured with errors, if the regressors are endogenous, or if there is unobserved cross-section dependence, the least squares estimates are inconsistent even when the sample sizes in both the time series (T) and the cross-section (N) dimensions are large. The source of these problems is the presence of a latent component in the model that is correlated with the regressors.

We provide a Panel-ME algorithm to show how dynamic panel models with measurement errors can be estimated without fully specifying the data generating processes for the covariates as in structural equation models of the LISREL type discussed in Joreskog and Thillo (1972). A simple modification leads to a Panel-IV algorithm that enables estimation of models with endogenous regressors without the need for instrumental variables. We also provide a Panel-CS algorithm for estimating panel models with cross-section dependence. Such models have generated a good deal of interest in recent years. Cross-section dependence in the form of a factor structure bears similarity with measurement errors in that they both involve a time varying latent component.

The appeal of IDEA lies in its applicability to a broad range of models and data configurations. Furthermore, it has a built in bias-correction property that is especially appealing when there is unobserved heterogeneity and T is small. This property was illustrated in Gourieroux, Phillips, and Yu (2010) using a panel AR(1) model. We show that IDEA shares this property even for models in which consistent estimation by the fixed effects estimator is not possible regardless of how large T is. In a companion paper Komunjer and Ng (2010), we show that IDEA also has excellent properties in small N and large T VARX models with measurement errors. We begin with an overview of the method of indirect inference of panel models with correctly observed data.

2 The Indirect Inference Estimator: An Overview

Let θ be the parameter vector in a model that may be complex, but can be easily simulated. Indirect inference requires specifying an auxiliary model that captures features of the data but need not coincide with the true model. Let $\vec{y}_i \equiv (y_{i1}, \ldots, y_{iT})'$ and $\vec{x}_i \equiv (x_{i1}, \ldots, x_{iT})'$ be the time series observations, and let $\mathbf{y} = (\vec{y}_1, \ldots, \vec{y}_N)$ and $\mathbf{x} = (\vec{x}_1, \ldots, \vec{x}_N)$ be the matrix of stacked up observations in the panel. Let then ψ be the parameters of the auxiliary model whose estimates are defined by

$$\widehat{\psi}_{N,T} = \text{argmax}_\psi Q_{N,T}(\mathbf{y}, \mathbf{x}, \psi; \theta_0).$$

Assuming it exists and is unique, the pseudo true value of ψ is given by $\psi^0_{\infty,T} = \text{argmax}_\psi \lim_{N\to\infty} Q_{N,T}(\mathbf{y}, \mathbf{x}, \psi; \theta_0)$. Note that both the estimator and the pseudo-true value of ψ depend on the objective function, Q.

Let $\mathbf{y}^s = (\vec{y}^s_1, \ldots, \vec{y}^s_N)$ with $\vec{y}^s_i = (y^s_{i1}, \ldots, y^s_{iT})'$ be data for the dependent variable simulated under the assumed true model, holding the exogenous covariates $\mathbf{x}$ fixed. This requires drawing u^s_{it} from a parametric distribution. Estimating the auxiliary model on the simulated data yields

$$\widetilde{\psi}^s_{N,T}(\theta) = \text{argmax}_\psi Q_{N,T}(\mathbf{y}^s, \mathbf{x}, \psi; \theta).$$

The IDE is obtained by solving

$$\widetilde{\theta}_{N,T,S} \equiv \text{argmin}_\theta \left\| \widehat{\psi}_{N,T} - \frac{1}{S}\sum_{s=1}^{S} \widetilde{\psi}^s_{N,T}(\theta) \right\|_{W_{N,T}},$$

where $W_{N,T}$ is a weighting matrix. Essentially, the auxiliary parameter estimates define a mapping from the parameter space of θ to the parameter space of the auxiliary model. Gourieroux, Monfort, and Renault (1993) refer to this mapping as a binding function.

Definition 1 *Let $\Psi : \theta_0 \to \Psi(\theta_0)$ be a mapping from θ_0 to $\psi^0_{\infty,T}$. Then θ_0 is globally and locally identified if $\Psi(\cdot)$ is injective and $\frac{\partial \Psi(\theta_0)}{\partial \theta}$ has full column rank.*

Although the auxiliary model need not nest the true model, it must contain features of the data generated under θ. In a likelihood setting, identification requires that the true densities of the data be 'smoothly embedded' within the scores of the auxiliary model, see Gallant and Tauchen (1996). When $Q_{N,T}(\cdot)$ is not likelihood based, identification requires that the conditional moments of the auxiliary model under θ have independent information about θ. This in turn means that ψ must be chosen such that $\frac{\partial \Psi(\theta)}{\partial \theta}$ has full column rank.

If the binding function were known, invertible, and θ of the same dimension as ψ, then $\widehat{\theta}_{N,T} = \Psi^{-1}(\widehat{\psi}_{N,T})$ would be a consistent estimate of θ_0. But $\Psi(\cdot)$ is usually an intractable function. The IDE uses simulations to approximate and invert the binding function. Under regularity conditions,

$$\sqrt{NT}(\widetilde{\theta}_{N,T,S} - \theta_0) \xrightarrow{d} N(0, \text{Avar}(\widetilde{\theta}_{N,T,S})),$$

where $\text{Avar}(\widetilde{\theta}_{N,T,S})$ is of the double sandwich form as defined in Proposition 3 of Gourieroux, Monfort, and Renault (1993). It depends on whether the auxiliary model is correctly specified via ψ, the asymptotic variance of $\widehat{\psi}_{N,T}$ through the choice of the Q, and the number of simulations, S.

In a recent paper, Gourieroux, Phillips, and Yu (2010) use the IDE to improve the estimates of a panel AR(1) model. Use of the IDE in dynamic panels is natural as Gourieroux, Renault, and Touzi (2000) show that the estimator automatically provides second order bias correction if the auxiliary model admits an Edgeworth expansion. Before explaining why the IDE will not work for measurement error models and explain what modifications are necessary, we first take a closer look at the IDE in the context of a panel ARX(1,0) model.

2.1 Direct and Indirect Bias Corrections: Case of ARX(1,0)

Consider a correctly measured dynamic panel model with fixed effects,

$$y_{it} = \lambda_i + A_1 y_{i,t-1} + B_0 x_{it} + u_{it}, \qquad i = 1, \ldots, N, t = 1, \ldots, T, \tag{1}$$

with $u_{it} \sim WN(0, \sigma_u^2)$. The parameters of this model are $\theta = (\theta^{+\prime}, \sigma_u^2)' = (B_0, A_1, \sigma_u^2)'$.[1] Let $\ddot{z}_{it} = (\ddot{y}_{i,t-1}, \ddot{x}_{it})'$ represent the deviation of $z_{it} = (y_{i,t-1}, x_{it})'$ from its time average. Let $\widehat{\sigma}_{\ddot{a}\ddot{b}}$ be the sample covariance between two demeaned variables, a and b. For example, $\widehat{\sigma}_{\ddot{x}\ddot{y}_{-1}} = \frac{1}{N(T-1)} \sum_{i=1}^{N} \sum_{t=1}^{T} \ddot{x}_{it} \ddot{y}_{i,t-1}$. For given N and T, the LSDV estimator $\widehat{\theta}_{N,T} = (\widehat{\theta}_{N,T}^{+\prime}, \widehat{\sigma}^2_{u,N,T})'$ is

$$\widehat{\theta}^+_{N,T} = \begin{pmatrix} \widehat{\sigma}_{\ddot{x}\ddot{x}} & \widehat{\sigma}_{\ddot{x}\ddot{y}_{-1}} \\ \widehat{\sigma}_{\ddot{x}\ddot{y}_{-1}} & \widehat{\sigma}_{\ddot{y}_{-1}\ddot{y}_{-1}} \end{pmatrix}^{-1} \begin{pmatrix} \widehat{\sigma}_{\ddot{x}\ddot{y}} \\ \widehat{\sigma}_{\ddot{y}_{-1}\ddot{y}} \end{pmatrix} \quad \text{and} \quad \widehat{\sigma}^2_{u,N,T} = \frac{1}{N(T-1)} \sum_{i=1}^{N} \sum_{t=1}^{T} (\ddot{y}_{it} - \ddot{z}_{it}' \widehat{\theta}^+_{N,T})^2.$$

We may write

$$\widehat{\theta}_{N,T} = \theta + b_{N,T}\Big(\widehat{\sigma}_{\ddot{y}_{-1}\ddot{y}_{-1}}, \widehat{\sigma}_{\ddot{x}\ddot{y}_{-1}}, \widehat{\sigma}_{\ddot{x}\ddot{x}}, \widehat{\sigma}_{\ddot{x}\ddot{u}}, \widehat{\sigma}_{\ddot{y}_{-1}\ddot{u}}\Big) \tag{2}$$

where $b_{N,T}(\cdot)$ is the bias function. Nickell (1981) showed that $\widehat{\theta}_{N,T}$ is inconsistent as $N \to \infty$ with T fixed. Consistent estimates can be obtained by various instrumental variable methods, details

[1] Our results are invariant to the assumptions made on the initial conditions. In the simulation experiments to follow, we shall assume the initial observation to be drawn as $y_{i0} \sim N\left(\frac{\lambda_i}{1-A_1} + B_0 x_{i0}, \frac{\sigma_u^2}{\sqrt{1-A_1^2}}\right)$.

of which can be found in Arellano (2003). However, the IV estimates are generally inefficient, and weak instruments pose a problem when the panel has near unit roots with A_1 close to unity.

An alternative to IV estimation is to obtain analytical approximations to the bias and then construct bias corrected estimators. Kiviet (1995), Bun and Carree (2005) and Phillips and Sul (2007) approximate the biases using large N asymptotics. Specifically, let $\sigma_{\ddot{a}\ddot{b}} \equiv \lim_{N\to\infty} \widehat{\sigma}_{\ddot{a}\ddot{b}}$. Then, as N gets large the bias function $b_{N,T}$ converges to $b_{\infty,T}(\theta, \sigma_{\ddot{y}_{-1}\ddot{y}_{-1}}, \sigma_{\ddot{x}\ddot{y}_{-1}}, \sigma_{\ddot{x}\ddot{x}})$; for analytic expressions of $b_{\infty,T}$ see, for example, Nickell (1981), Kiviet (1995), and Bun and Carree (2005). The bias-corrected estimator $\widehat{\theta}_{BC}$ is then defined as the value of θ that solves

$$\widehat{\theta}_{N,T} = \theta + b_{\infty,T}\bigg(\theta, \widehat{\sigma}_{\ddot{y}_{-1}\ddot{y}_{-1}}, \widehat{\sigma}_{\ddot{x}\ddot{y}_{-1}}, \widehat{\sigma}_{\ddot{x}\ddot{x}}\bigg). \tag{3}$$

Large N and T approximations has been used by Hahn and Kuersteiner (2002) to obtain $b_{\infty,\infty}(\cdot)$. Kiviet (1995) showed that these bias-corrected estimators tend to have smaller root mean-squared error (RMSE) than the IV estimators. However, the bias corrections are model specific and are invalid when there are additional covariates or lagged independent variables.

Consider now the IDE based on simulated u^s_{it}, y^s_{it} and observed x_{it}. Let the auxiliary model be the (true) ARX(1,0) model and consider estimating the model by LSDV. Under these assumptions, $\theta = \psi$ and the choice of $W_{N,T}$ does not matter. Then

$$\widetilde{\theta}^s_{N,T} \;=\; \theta + b^s_{N,T}\bigg(\widehat{\sigma}_{\ddot{y}^s_{-1}\ddot{y}^s_{-1}}(\theta), \widehat{\sigma}_{\ddot{x}\ddot{y}^s_{-1}}(\theta), \widehat{\sigma}_{\ddot{x}\ddot{x}}, \widehat{\sigma}_{\ddot{x}\ddot{u}^s}(\theta), \widehat{\sigma}_{\ddot{y}^s_{-1}\ddot{u}^s}(\theta)\bigg). \tag{4}$$

Note that unlike in (2), the covariance terms obtained from simulated u^s_{it} and y^s_{it} now explicitly depend on θ. The IDE is then obtained as a solution $\widetilde{\theta}_{N,T,S}$ to

$$\min_\theta \left\|\widehat{\theta}_{N,T} - \frac{1}{S}\sum_{s=1}^{S}\widetilde{\theta}^s_{N,T}\right\| = \tag{5}$$

$$\min_\theta \left\| b_{N,T}\bigg(\widehat{\sigma}_{\ddot{y}_{-1}\ddot{y}_{-1}}, \widehat{\sigma}_{\ddot{x}\ddot{y}_{-1}}, \widehat{\sigma}_{\ddot{x}\ddot{x}}, \widehat{\sigma}_{\ddot{x}\ddot{u}}, \widehat{\sigma}_{\ddot{y}_{-1}\ddot{u}}\bigg) - \frac{1}{S}\sum_{s=1}^{S} b^s_{N,T}\bigg(\widehat{\sigma}_{\ddot{y}^s_{-1}\ddot{y}^s_{-1}}(\theta), \widehat{\sigma}_{\ddot{x}\ddot{y}^s_{-1}}(\theta), \widehat{\sigma}_{\ddot{x}\ddot{x}}, \widehat{\sigma}_{\ddot{x}\ddot{u}^s}(\theta), \widehat{\sigma}_{\ddot{y}^s_{-1}\ddot{u}^s}(\theta)\bigg)\right\|$$

The IDE exploits the fact that if the LSDV estimator $\widehat{\theta}_{N,T}$ based on the true data is biased, then $\widetilde{\theta}^s_{N,T}$ which is based on the simulated data will also be biased. Calibrating the bias terms by simulations eliminates the need for analytical derivations on a model specific basis.

As noted above, the IDE depends on the simulated variables y^s_{it} and u^s_{it} which depend on θ; hence, all the covariance terms appearing in the simulated bias $b^s_{N,T}$, except $\widehat{\sigma}_{\ddot{x}\ddot{x}}$, are functions of θ. In contrast, all the covariance terms appearing in the limit bias $b_{\infty,N}$, and thus in $\widehat{\theta}_{BC}$ in (3), are invariant to θ. As the binding function of the two estimators vary with θ in different ways, they will have also different variance in finite samples, even though both are of the double sandwich form. However, as S and N tend to ∞, the right hand side of (4) evaluated at θ_0 converges to the right hand side of (3). Thus, the asymptotic variance of $\widetilde{\theta}_{N,T,S}$ is the same as $\widehat{\theta}_{BC}$ given in Bun and Carree (2005).

Notice that (4) reflects the fact that the covariates x_{it} are held fixed in the simulations. This is valid if the parameters of the conditional model do not vary with those in the marginal distribution of the covariates. Or in other words, x is weakly exogenous for θ in the sense of Engle, Hendry, and Richard (1983). While such exogeneity conditions are typically assumed to hold in correctly measured panel models, they break down as soon as measurement errors are present.

3 An Augmented Indirect Estimator: IDEA

As far as we are aware of, the only reference to IDE in the measurement error literature is Jiang and Turnbull (2004) who use indirect inference as a way to adjust the bias in the auxiliary parameters with the help of validation data. Standard implementation of the IDE without validation data is problematic because the regressors are no longer weakly exogenous for the parameters of the conditional model. More precisely, when the parameters in the conditional model and those in the marginal density of the covariates are no longer 'variation free', the covariates x_{it} cannot be held fixed in simulations. To make this point precise, consider the ARX(1,0) model without fixed effect but with measurement error in $X_{it} = x_{it} + \epsilon^x_{it}$:

$$
\begin{aligned}
y_{it} &= \lambda_i + A_1 y_{i,t-1} + B_0 x_{it} + u_{it} \\
&= \lambda_i + A_1 y_{i,t-1} + B_0 X_{it} + V_{it}.
\end{aligned}
$$

As $V_{it} = u_{it} - B_0 \epsilon^x_{it}$ is not orthogonal to X_{it}, an OLS estimator $\widehat{\psi}_{N,T}$ of the parameter $\psi = (B_0, A_1)$ based on the observed data is such that $\text{plim}_{N\to\infty} \widehat{\psi}_{N,T} = \psi + \text{bias} \neq \psi$. Now consider estimating the model by the IDE. If X_{it} is fixed in simulations, then $X_{it} \perp V^s_{it} = u^s_{it} - B_0 \epsilon^{sx}_{it}$ by construction, and plim $\widehat{\psi}^s_{N,T} = \psi$. The binding function $\Psi(\theta)$ will not be consistently estimated.

The next subsection shows how to define ψ for the purpose of identifying the parameters of interest in the presence of latent components, and how to simulate the covariates without fully specifying their data generating process. The key idea is to use the relation

$$
\text{var}(X_{it}) = \text{var}(x_{it}) + \text{var}(\epsilon^x_{it})
$$

to obtain estimates of the latent x_{it} from the data up to scale. Then X^s_{it} can be obtained upon simulation of ϵ^x_{it}.

3.1 The Case of Measurement Errors

The LSDV estimates of the ARX(1,0) model are biased even without measurement errors. However, with or without measurement error, $\widehat{\psi}_{N,T}$ is consistent for the pseudo parameter $\psi^0_{\infty,T}$ even though $\psi^0_{\infty,T} \neq \theta_0$. It remains to be precise what is θ and what is the ψ appropriate for the measurement error model. To this end, we consider a general panel autoregressive distributed lag ARX(p_y, r_x) model with dependent variable y_{it} and a scalar covariate x_{it} given by

$$y_{it} = \lambda_i + \sum_{\tau=1}^{p_y} A_\tau y_{i,t-\tau} + \sum_{\tau=0}^{r_x} B_\tau x_{i,t-\tau} + u_{it}, \qquad (6)$$

where $i = 1, \ldots, N$ and $t = 1, \ldots, T$. The observed data are possibly contaminated by classical additive errors:

$$\begin{aligned} Y_{it} &= y_{it} + \epsilon^y_{it}, \\ X_{it} &= x_{it} + \epsilon^x_{it}. \end{aligned}$$

Let $n^y_\epsilon = 1$ if y_{it} is measured with error and zero otherwise. Similarly, let $n^x_\epsilon = 1$ if the covariate x_{it} is measured with error and zero otherwise. We assume the measurement errors to be drawn from independent ARMA processes: $\phi^y(L)\epsilon^y_{it} = \vartheta^y(L)v^y_{it}$ and $\phi^x(L)\epsilon^x_{it} = \vartheta^x(L)v^x_{it}$ in which $v^y_{it} \sim WN(0, \sigma^{y2}_v)$ and $v^x_{it} \sim WN(0, \sigma^{x2}_v)$ are independent.

Let V_{it} be the errors of the auxiliary model let $\Gamma_V(j)$ be the (possibly vector of) autocovariance of V_{it} at lag j. With the auxiliary model being the true model, consider

$$\begin{aligned} \theta^y_\epsilon &= (\phi^y, \vartheta^y, \sigma^y_v) \quad \text{if } n^y_\epsilon = 1 \\ \theta^x_\epsilon &= (\phi^x, \vartheta^x, \sigma^x_v) \quad \text{if } n^x_\epsilon = 1 \\ \theta &= (A_1, \ldots, A_{p_y}, B_0, \ldots, B_{r_x}, \sigma^2_u, \theta^y_\epsilon, \theta^x_\epsilon) \\ \psi &= (A_1, \ldots, A_{p_y}, B_0, \ldots, B_{r_x}, \Gamma_V(0), \ldots, \Gamma_V(q_V)), \end{aligned}$$

As the parameters $(A_1, \ldots, A_{p_y}, B_0, \ldots, B_{r_x}, \sigma^2_u)$ cannot be separately identified from those of the measurement error processes, θ is augmented to include nuisance parameters belonging to the measurement errors. The auxiliary parameter vector ψ crucially depends on $\Gamma_V(q_V)$ for suitably chosen q_V.

Identification: Panel-ME (i) $\frac{\partial \Psi(\theta)}{\partial \theta}$ is full column rank; and (ii) $q_V \geq \dim(\theta^y_\epsilon) + \dim(\theta^x_\epsilon)$.

Part (i) is the necessary and sufficient rank condition for identification of θ; part (ii) is the necessary order condition. In the absence of measurement error, identification would have required the information matrix of the ARX model (which is a function of θ) to be full rank, as in Rothenberg (1971). Instead, our rank condition concerns the Jacobian matrix of the $\Psi(\theta)$ with respect to θ, which can be checked analytically or numerically.

The above identification analysis stands in sharp contrast to the conventional IV approach. With the latter, one transforms the model to remove the bias-inducing terms so that orthogonaltiy conditions can be found. While instrumental variables are required, parametric specification of the measurement error process or of the covariates is not necessary. On the other hand, IDEA exploits information in the bias. It does not require instruments, but demands specification of the measurement error process.

For the ARX(1,0) model with $\theta^+ = (B_0, A_1)'$, the indirect estimator is

$$\widetilde{\theta}^{s+}_{N,T} \quad = \quad \theta^+ + b^s_{N,T}\Big(\widehat{\sigma}_{\ddot{Y}^s_{-1}\ddot{Y}^s_{-1}}(\theta), \widehat{\sigma}_{\ddot{X}^s\ddot{Y}^s_{-1}}(\theta), \widehat{\sigma}_{\ddot{X}^s\ddot{X}^s}(\theta), \widehat{\sigma}_{\ddot{X}^s\ddot{V}^s}(\theta), \widehat{\sigma}_{\ddot{Y}^s_{-1}\ddot{V}^s}(\theta)\Big). \tag{7}$$

Unlike (4) when there is no measurement error, $\widehat{\sigma}_{\ddot{X}\ddot{X}}(\theta)$ in (7) is now a function of θ via the measurement error parameters. This also makes clear that simulation of the contaminated data X_{it} is now necessary. We propose the following:

Algorithm Panel-ME

1. Estimate the auxiliary model to yield $\widehat{\psi}^+_{N,T} = (\widehat{A}_1, \ldots, \widehat{A}_{p_y}, \widehat{B}_0, \ldots, \widehat{B}_{r_x})'$ and residuals $\widehat{V}_{i,t}$.

2. Compute its $T \times T$ variance covariance matrix with (t, j) element being $\Gamma_V(t, j) = \frac{1}{N}\sum_{i=1}^N \widehat{V}_{it}\widehat{V}_{i,t-j}$. Let $\Gamma_V(j) = \frac{1}{T-j}\sum_{t=j+1}^T \Gamma_V(t, j)$.

3. Given a guess of θ and for $s = 1, \ldots, S$:

   a. Draw $v^{ys}_{it} \sim N(0, \sigma^2_{v^y})$ to obtain $\phi^y(L)\epsilon^{ys}_{it} = \vartheta^y(L)v^{ys}_t$.

   b. Draw $v^{xs}_{it} \sim N(0, \sigma^2_{v^x})$ to obtain $\phi^x(L)\epsilon^{xs}_{it} = \vartheta^x(L)v^{xs}_{it}$.

   c. Compute $\text{var}(\epsilon^{xs}_{it})$ and let $x^s_{it} = X_{it} \cdot \left[1 - \frac{\text{var}(\epsilon^{xs}_{it})}{\text{var}(X_{it})}\right]^{1/2}$.

   d. Draw $u^s_{it} \sim N(0, \sigma^2_u)$. Construct $y^s_{it} = \sum_{j=1}^{p_y} A_j y^s_{i,t-j} + \sum_{j=0}^{r_x} B_j x^s_{it} + u^s_{it}$.

   e. Let $X^s_{it} = x^s_{it} + \epsilon^{xs}_{it}$ and $Y^s_{it} = y^s_{it} + \epsilon^{ys}_{it}$.

   f. Estimate the auxiliary model using $(\mathbf{Y}^s, \mathbf{X}^s)$ to obtain $\widetilde{\psi}^{s+}_{N,T}$. Construct the autocovariances for $\widetilde{V}^s_{it}$ as in step 2. Then $\widetilde{\psi}^s_{N,T} = (\widetilde{\psi}^{s+}_{N,T}, \Gamma_{\widetilde{V}}(q_V))$.

As is well documented, the choice of auxiliary model and estimator affect the efficiency of the IDE. Li (2010) favors OLS estimation of a linear auxiliary model for estimating structural auction models with computational ease being one of the reasons. Michaelides and Ng (2000) find that an auxiliary model with thresholds yields more precise estimates than a linear model . Here, we let the auxiliary model be the true model. Still, the parameters of the auxiliary model can be estimated in Step f by multiple ways. The obvious choice is LSDV since it is efficient when $N, T \to \infty$. A different implementation is to estimation ψ by instrumental variables. OLS estimation is also possible as we do not need $\widehat{\psi}_{N,T}$ to be consistent for θ_0.

3.2 The Case of Endogeneity

It is known that the coefficients associated with endogenous regressors cannot be consistently estimated when the variations in the dependent variable can no longer be traced to exogenous variations in the regressors. As with the problem of measurement error, the orthogonality between the regression error and the regressor breaks down. Consider the stationary dynamic panel model

$$A(L)y_{it} = \lambda_i + B(L)X_{it} + u_{it}$$

where u_{it} is serially uncorrelated, but the scalar covariate X_{it} is contemporaneously correlated with u_{it}. Specifically, assume that

$$X_{it} = \gamma u_{it} + x_{it}$$

has two mutually orthogonal latent components: an x_{it} satisfying $E(x_{it}u_{it}) = 0$, and a component that induces endogeneity. If x_{it} were observed, it would have been the ideal instrument as it is correlated with X_{it} and uncorrelated with u_{it}. The measurement error problem is evidently a special case of endogeneity with $\gamma = 1$. To identify the exogenous shifts in X_{it} as given by x_{it} on y_{it}, we will exploit the fact that $\text{var}(X_{it}) = \gamma^2\sigma_u^2 + \text{var}(x_{it})$. Define

$$\begin{aligned}\theta &= (A_1, \ldots, A_{p_y}, B_0, \ldots B_{r_x}, \sigma_u^2, \gamma)\\ \psi &= (A_1, \ldots, A_{p_y}, B_0, \ldots, B_{r_x}, \Gamma_u(0), \ldots, \Gamma_u(q_u)).\end{aligned}$$

Algorithm Panel-IV

1. Estimate the auxiliary model to yield $\widehat{\psi}^+_{N,T} = (\widehat{A}_1, \ldots, \widehat{A}_{p_y}, \widehat{B}_0, \ldots, \widehat{B}_{r_x})'$ and residuals $\widehat{u}_{it}$.

2. Compute its $T \times T$ variance covariance matrix with (t, j) element being $\Gamma_u(t, j) = \frac{1}{N}\sum_{i=1}^N \widehat{u}_{it}\widehat{u}_{it-j}$. Let $\Gamma_u(j) = \frac{1}{T-j}\sum_{t=j+1}^T \Gamma_u(t, j)$.

3. Given a guess of θ and for $s = 1, \ldots, S$:

   a. Let $x^s_{it} = X_{it}\left[1 - \frac{\gamma^2\sigma_u^2}{var(X_{it})}\right]^{1/2}$. Draw $u^s_{it} \sim N(0, \sigma_u^2)$ and let $X^s_{it} = \gamma u^s_{it} + x^s_{it}$.

   b. Construct $y^s_{it} = \sum_{j=1}^{p_y} A_j y^s_{i,t-j} + \sum_{j=0}^{r_x} B_j x^s_{it} + u^s_{it}$.

c. Estimate the auxiliary model using $(\mathbf{y}^s, \mathbf{X}^s)$ to obtain $\widetilde{\psi}^{s+}_{N,T}$. Construct the autocovariances for $\widetilde{u}^s_{it}$ as in step 2. Then $\widetilde{\psi}^s_{N,T} = (\widetilde{\psi}^{s+}_{N,T}, \Gamma_{\widetilde{u}}(q_u))$.

Algorithm Panel-IV presents an alternative solution to the endogeneity problem which usually consists of finding instruments that are exogenous yet strongly correlated with X_{it}. Parametric assumptions about the nature of endogeneity are not required. Instead, IDEA takes a control function type approach to make explicit the relation between the equation error and X_{it}. However, the need for instruments is dispensed at the cost of having to specify how X_{it} is correlated with u_{it}. Our linearity assumption leads to a simple way that allows the exogenous variations in X_{it} to be simulated.

3.3 The Case of Cross-Section Dependence

A generalization of the fixed effects model is a multiplicative two-way model that allows the unobserved heterogeneity to have time varying effects. The model is

$$Y_{it} = \lambda_i^{y\prime} f_t + A_1 Y_{i,t-1} + X_{it}' B_0 + u_{it} \quad (8)$$

$$X_{it} = \lambda_i^{x\prime} g_t + x_{it} \quad (9)$$

where λ_i^y and f_t are $r_f \times 1$ vectors, while λ_i^x and g_t are $r_g \times 1$.[2] In this model, λ_i^y and λ_i^x are unobserved (vectors of) individual effects whose impact on Y_{it} and X_{it} may change over time according to f_t and g_t respectively. Assuming $g_t = 0$ for all t, Kiefer (1980) derives a concentrated least squares estimator for the $r = 1$ case, while Holtz-Eakin, Newey, and Rosen (1988) suggest a quasi-differencing approach to estimate panel vector autoregressive models in the presence of $\lambda_i^{y\prime} f_t$. Ahn, Lee, and Schmidt (2001) exploit orthogonaltiy conditions to estimate the model by GMM, while Nauges and Thomas (2003) use a double difference transformation to estimate the model, also by GMM. Bai (2009) considers maximum likelihood estimation treating f_t and g_t as parameters using the approach of Chamberlain (1982) and Mindlak (1978) to control for the correlation between the fixed effects and the regressors. The likelihood depends on whether or not the regressors are weakly exogenous.

We consider indirect estimation of A_1 and B_0 simultaneously with f_t and g_t treated as parameters. This is motivated by the observation that the latent multiplicative fixed effects $\lambda_i^{y\prime} f_t$ and $\lambda_i^{x\prime} g_t$ in (8) and (9) are time varying latent components, much like measurement errors. Whereas parametric assumptions were made for the measurement error process, we now need to be precise about λ_i^x and λ_i^y.

[2] As pointed out before, our approach allows for arbitrary assumptions regarding the initial conditions. In the simulations carried out in the next section, we shall assume that $Y_{i0} \sim N\left(\frac{\lambda_i^{y\prime} f_0}{1-A_1} + X_{i0}' B_0 \frac{\sigma_u^2}{\sqrt{1-A_1^2}}\right)$.

Let the auxiliary model be the true model (8) with residuals u_{it}. Define

$$\begin{aligned}\theta &= \left(A_1, B_0, \sigma_u, \bar{\lambda}^y, \bar{\lambda}^x, \sigma_{\lambda^y}, \sigma_{\lambda^x}, \{f_t\}_{t=1}^T, \{g_t\}_{t=1}^T\right)' \\ \psi &= \left(A_1, B_0, \vec{\Gamma}_u(q_u)\right)',\end{aligned}$$

where $\vec{\Gamma}_u(q_u)$ is a $q_u \times 1$ sub-vector of the unique elements of the $T \times T$ covariance matrix of u_{it}.

Identification: Panel-CS: (i) $\frac{\partial \Psi(\theta)}{\partial \theta}$ is full column rank; and (ii) $q_u \geq (T+2)(r_g + r_f)$.

The rank condition requires that the auxiliary model encompasses features of the model of interest. As θ is a subvector of ψ, the condition reduces to $\frac{\partial \Gamma_u(\theta)}{\partial \theta}$ being full rank. The necessary (order) condition is, however quite stringent as we now have to estimate T elements of the r_f vector f_t and T elements of the r_g vector g_t. As the $T \times T$ covariance matrix of u_{it} can have at most the $T(T+1)/2$ elements, the order condition requires that $T \geq 5$.

Algorithm Panel-CS

1 Choose a criterion function Q to obtain $\widehat{\psi}^+_{N,T} = (\widehat{A}_1, \widehat{B}_0)$ and residuals $\widehat{u}_{it}$.

2. Construct $\Gamma_u(t,j) = \frac{1}{N}\sum_{i=1}^N \widehat{u}_{it}\widehat{u}_{i,t-j}$. Let

$$\vec{\Gamma}_u(q_u) = (\{\Gamma_u(t,0)\}_{t=1}^T, \{\Gamma_u(t,1)\}_{t=2}^T, \ldots \{\Gamma_u(t,k)\}_{t=k+1}^T\}$$

where k is chosen to satisfy the order condition.

3 Given a guess of θ, repeat for $s = 1, \ldots, S$:

   a. Simulate the $r_f \times 1$ vector λ_i^{ys} and the $r_g \times 1$ vector λ_i^{xs} from assumed distributions;

   b. Let $X_{it}^s = x_{it}^s + \lambda_i^{xs\prime} g_t$ with

$$x_{it}^s = X_{it} \cdot \left[1 - \frac{g_t' \sigma^2_{\lambda^x} g_t}{\text{var}(X_{it})}\right]^{1/2}.$$

   c. Draw u_{it}^s and let $Y_{it}^s = \lambda_i^{ys\prime} f_t + A_1 Y_{i,t-1}^s + X_{it}^{s\prime} B_0 + u_{it}^s$.

4. Estimate the auxiliary model on $(\mathbf{Y}^s, \mathbf{X}^s)$ according to criterion $Q(\cdot)$ to yield $\psi^{+s}_{N,T}$. Construct $\widetilde{\Gamma}_u^s(t,j)$ as in step 2.

In the stationarity measurement error model, we average $\Gamma_V(t,j)$ over t to obtain $\Gamma_V(j)$. With Algorithm MFE, $\Gamma_u(t,j)$ at each t is a distinct entry of $\widetilde{\psi}_{N,T}$. Thus if $T=5$, there would be five $\Gamma_u(t,0)$, four $\Gamma_u(t,1)$ and so forth. This is necessary to identify f_t and g_t for $t=1,\ldots T$.

Two additional implementation issues need to be highlighted. First, in the measurement error model, demeaning removes λ_i. Thus any value of λ_i (including zero) can be used to simulate the data. The LSDV/ FD estimates of the auxiliary parameters are invariant to what is assumed for λ_i. In the present setup, demeaning and differencing no longer remove λ_i. For this reason, the first step in each simulation is to make draws of λ_i^y and λ_i^x. Second, so long as f_t and g_t are time varying, LSDV will not be consistent even when N and T are both large. Although LSDV can still be used, it is no longer the obvious candidate estimator for the parameters of the auxiliary (also the true) model.

4 Simulations

This section consists of four parts. Subsection 1 assesses the finite sample properties of the IDE for the dynamic panel model with no measurement error. Subsection 2 considers ARX(1,1) models with measurement errors. Endogeneity is considered in subsection 3. Subsection 4 considers the case cross-section dependence. The mean and root-mean-squared error of the estimates are computed from 1000 replications. The IDE and IDEA results are based on $S=100$ draws.

4.1 The ARX(1,0) Model

Data are generated according to (1):

$$y_{it}=\lambda_i+A_1y_{i,t-1}+B_0x_{it}+u_{it}$$

where for $t\geq 1$, $u_{it}\sim\sigma_{u0}N(0,.95-05T+.1t)$. Furthermore, $\lambda_i\sim N(0,\sigma_\lambda^2)$ and $y_{i0}\sim N(\frac{\lambda_i}{1-A_1}+B_0x_{i0},\frac{\sigma_{u0}^2}{\sqrt{1-A_1^2}})$. The scalar covariate x_{it} is generated as

$$x_{it}=\mu^x+A_1^x x_{i,t-1}+u_{it}^x,\qquad x_{i0}\sim N\left(\frac{\mu^x}{1-A_1^x},\frac{\sigma_v^2}{\sqrt{1-(A_1^x)^2}}\right),\qquad u_{it}^x\sim N(0,\sigma_x^2). \tag{10}$$

We let $A_1^x=.8$, $\mu^x=0,\sigma_x=1$. The true values of (B_0,A_1,σ_{u0}) are (1, .8, 1). We consider the six parameter configurations in Tables 2-4 of Bun and Carree (2005), For the sake of comparison, we also compute the LSDV, along with the bias-corrected estimator of Bun and Carree (2005).[3] Note that the fixed effect is absent in x_{it}.

[3]As Bun and Carree (2005) also consider the GMM estimator of Arellano and Bond (1991) and the modified corrected estimator of Kiviet (1995), these results will not be reported. It should however be mentioned that $\widehat{\theta}_{bc}$ has smaller bias than the other estimators considered.

The results are reported in Table 1. The LSDV estimates of A_1 are strongly biased downwards when $T = 2$ or 3, similar to the panel AR(1) results reported in Gourieroux, Phillips, and Yu (2010) for $T = 6$. The $\widehat{B}_{0,BC}$ and $\widehat{B}_{0,IDE}$ are much more precise. The IDE tends to provide somewhat more efficient estimates of A_1 when T is small. Results with $u_{it} \sim N(0, \sigma_u^2)$ and cross-section heteroskedasticity with $\sigma_{ui}^2 \sim \chi_1^2$ are similar and not reported to conserve space. Suffice it to mention that $\widehat{\theta}_{IDE}$ has better properties when $T < 6$ (especially under cross-section heteroskedasticity) while $\widehat{\theta}_{BC}$ is better with larger T. Overall, $\widehat{\theta}_{BC}$ and $\widehat{\theta}_{IDE}$ have similar finite sample properties. The advantage of the IDE is that it can be used in models that are not specified exactly as in (1), such as when there are multiple covariates or additional lags.

4.2 The ARX(1,1) Model with Measurement Errors

Data are generated as follows:

$$
\begin{aligned}
y_{it} &= \lambda_i + A_1 y_{i,t-1} + B_0 x_{it} + B_1 x_{it-1} + u_{it} \\
x_{it} &= \lambda_i + \rho x_{i,t-1} + u_{it}^x \\
X_{it} &= x_{it} + \epsilon_{it}^x, \quad \epsilon_{it}^x = \phi_x \epsilon_{i,t-1} + v_{it}, \quad v_{it}^x \sim N(0, \sigma_{vx}^2) \\
Y_{it} &= y_{it} + \epsilon_{it}^y, \quad \epsilon_{it}^y = \phi_x \epsilon_{i,t-1}^y + v_{it}^y, \quad v_{it}^y \sim N(0, \sigma_{vy}^2)
\end{aligned}
$$

with $\lambda_i \sim N(0,1)$, $\sigma_u^2 = 1$. The covariate x_{it} now has a fixed effect. We let $\rho = 0.5$. Simulations are performed with $N = 200$ and $q_V = 2$. Thus, $\dim(\theta) = \dim(\psi) = 6$.

We first verify that the IDEA is not affected by having to estimate the measurement error parameters unnecessarily. This is reported in Table 2. The infeasible LSDV using the correctly observed data $(\mathbf{y}, \mathbf{x})$ are reported in column 1. The IDEA estimates are based on using the true model as the auxiliary model. Two estimators of the auxiliary parameters are considered: OLS (column 2) and LSDV (column 3). The IDEA estimates using OLS are inferior to those based on LSDV and it is useful to understand why. As demeaning removes the fixed effect, LSDV estimates are invariant to λ_i. The mean and variance of λ_i can conveniently be set to zero. With OLS, the mean of λ_i is absorbed in the intercept leaving $\lambda_i - \bar{\lambda}$ in the residuals. Even if θ is augmented to include σ_λ^2, we cannot separately identify σ_λ from the measurement error parameters. In other words, $\frac{\partial \Psi(\theta)}{\partial \theta}$ fails to be full rank. The highlights the point that the choice of the auxiliary model and the estimator for its parameters are both important. Indirect estimation will yield consistent and improved estimates only when identification conditions are met. The OLS based IDEA estimates will not be reported in subsequent results for the fixed effect model.

We then add AR(1) measurement errors to either y_{it} or x_{it} or both. Instead of the OLS based IDEA estimates, we now report the LSDV using contaminated data $(\mathbf{Y}, \mathbf{X})$ to gauge the extent of measurement error bias. These results are reported in Tables 3a and 3b. In all four cases, $\dim(\psi) = \dim(\theta)$. The IDEA estimates are generally precise when the data are truly observed with serially correlated AR(1) errors. The estimates of measurement error parameters exhibit downward bias when $T = 5$, but are precise when $T = 10$. Results for A_1 and B_0 are also similar if λ_i is drawn from the uniform distribution but a normal distribution is assumed. Varying σ_v relative to σ_u has little change on the results.

4.3 Endogeneity

Data are generated as follows:

$$\begin{aligned} y_{it} &= \lambda_i + A_1 y_{i,t-1} + B_0 X_{it} + u_{it} \\ x_{it} &= \lambda_i + \rho x_{i,t-1} + u^x_{it}, \\ X_{it} &= x_{it} + \gamma u_{it} \end{aligned}$$

where $u^x_{it} \sim N(0,1)$ and independent of $u_{it} \sim N(0,1)$. In the simulations, we let $\gamma = .5$. The results for $N = 200$ and $T = 5, 10$ are reported in Table 4. Also reported are the infeasible LSDV estimates using x_{it} as regressor (denoted LSDV$_x$), and the LSDV estimates using the endogenous regressor X_{it} (denoted LSDV$_X$).

When $\gamma = 0$ and hence there is no endogeneity, all three estimators have precise estimates of B_0 though IDEA is less efficient. However, IDEA has a more precise estimate of A_1. When X_{it} is endogenous with $\gamma = .5$, the infeasible LSDV$_x$ is still precise for B_0, but LSDV$_X$ is strongly upward biased. The bias persist even when $T = 10$. In contrast, the IDEA estimate of both B_0 and A_1 are precise. Furthermore, the variance of the estimates falls as T increases.

4.4 Cross-Section Dependence

We simulate data according to (8) and (9) with $\lambda^y_i = \lambda^x_i \sim U(.5, 1)$ for all i. We consider two serially correlated regressors. For $k = 1, 2$,

$$\begin{aligned} f_t &= \rho^f f_{t-1} + u^f_t, \qquad u^f_t \sim N(0, \sigma^2_f) \\ X_{ikt} &= \lambda_i g_t + \mu_k + x_{ikt}, \quad g_t \sim N(0,1), \\ x_{ikt} &= A^x_{1k} x_{ik,t-1} + u^x_{it}, \qquad u^x_{ikt} \sim N(0,1). \end{aligned}$$

with $\rho^f = A^x_1 = 0.5$, $\mu_1 = .25, \mu_2 = .5$. We let $(N, T) = (200, 5)$. The initial condition is

$$Y_{i0} \sim N\left(\frac{\lambda^{y\prime}_i f_0}{1 - A_1} + X'_{i0} B_0, \frac{\sigma^2_u}{\sqrt{1 - A^2_1}}\right),$$

where f_0 and X_0 are draws from their unconditional distributions.

In the simulations, we use $q_u = 2$, giving $\dim(\theta) = 16$ and $\dim(\psi) = 18$. The IDEA estimates are based on OLS estimation of the auxiliary model, which coincides with the true model.

When f_t is time varying, demeaning no longer removes the fixed effect. Both LSDV and OLS yield bised estimates even when N is large. We use OLS to estimate the auxiliary parameters as they more precise than than the LSDV estimates. The results are reported in Table 5. The first panel sets $\sigma_f^2 = 0$ to verify the IDEA is unaffected by having to estimate f_t and g_t unnecessarily. Indeed, in such a case, OLS, LSDV, and IDEA are equally precise. Panels two and three sets allow for a non-degenerate multiplicative fixed effect but u_{it} is homoskedastic. Evidently, all three estimators are biased, but the IDEA produces much more precise estimates of A_1 and B_0. With $T = 5$, these results are surprisingly good. Introducing time series heteroskedasticity as in Table 5b not change the picture. In fact, IDEA appears to be more precise under heteroskedasticity.

5 Conclusion

This paper considers an indirect estimation method for estimating models with time varying latent effects. The augmented indirect estimator (IDEA) is used to estimate dynamic panel models with measurement errors, endogenous regressors, and cross section dependence. The estimator has a built in bias-correction property that can be especially useful when T is small. The finite sample properties of IDEA are promising. In particular, the estimator suffers only small efficiency loss when there truly is no measurement error, endogeneity, or cross-section dependence. IDEA exploits information in the biases rather than removes the biases from orthogonality conditions. While IDEA requires the practitioner to specify how the regressors are correlated with the error term, full specification of the data generating process of the regressors is not necessary. The appeal of IDEA is generality and its ability to automatically perform bias correction when T is small.

References

Ahn, S., H. Lee, and P. Schmidt (2001): "GMM Estimation of Linear Panel Data Models with Time Varying Individual Effects," *Journal of Econometrics*, 101, 219–255.

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

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, 277–297.

Bai, J. (2009): "Likelihood Approach to Small T Dynamic Panel Models with Interactive Effects," mimeo, Columbia University.

Bun, M. J., and M. A. Carree (2005): "Bias-Corrected Estimation in Dynamic Panel Data Models," *Journal of Business and Economic Statistics*, 23, 200–210.

Chamberlain, G. (1982): "Multivariate Regression Models for Panel Data," *Journal of Econometrics*, 18, 5–46.

Engle, R., D. Hendry, and J. Richard (1983): "Exogeneity," *Econometrica*, 51:2, 277–304.

Gallant, R., and G. Tauchen (1996): "Which Moments to Match," *Econometric Theory*, 12, 657–681.

Gourieroux, C., A. Monfort, and E. Renault (1993): "Indirect Inference," *Journal of Applied Econometrics*, 85, 85–118.

Gourieroux, C., E. Renault, and N. Touzi (2000): "Calibration by Simulation for Small Sample Bias Correction," in *Simulation Based Inference in Econometrics: Methods and Applications*, ed. by R. Mariano, T. Schuermann, and M. Weeks, pp. 328–358.

Gourieroux, G., P. Phillips, and Y. Yu (2010): "Indirect Inference of Dynamic Panel Models," *Journal of Econometrics*.

Hahn, J., and G. Kuersteiner (2002): "Asymptotically Unbiased Inference for a Dynamic Model with Fixed Effects when both n and T are large," *Econometrica*, 70, 1639–1657.

Holtz-Eakin, D., W. Newey, and H. Rosen (1988): "Estimating Vector-Autoregrsesions with Panel Data," *Econometrica*, 56, 1371–1395.

Jiang, W., and B. Turnbull (2004): "The Indirect Method: Inference Based on Intermediate Statistics: A Synthesis and Examples," *Statistical Science*, 19:2, 239–263.

Joreskog, K., and V. Thillo (1972): *LISREL: A General Computer Program for Estimating a Linear Sturucal Eqauation System Invilving Multiple Indicators of Unmeasureed Variables*. Princeton, N.J.

Kiefer, N. (1980): "A Time Series Cross-Section Model with Fixed Effects with an Intertemporal Factor Structure," Cornell University.

Kiviet, J. (1995): "On Bias, Inconsistency and Efficiency of Various Estimators in Dynamic Panel Data Models," *Journal of Econometrics*, 68, 53–78.

Komunjer, I., and S. Ng (2010): "Measurement Errors in Dynamic Models," unpublished manuscript, Columbia University.

Li, T. (2010): "Indirect Inference in Structural Econometric Models," *Journal of Econometrics*, 157:1, 120–128.

Michaelides, A., and S. Ng (2000): "Estimating the Rational Expectations Model of Speculative Storage: A Monte Carlo Comparison of Three Simulation Estimators," *Journal of Econometrics*, 96:2, 231–266.

Mindlak, Y. (1978): "On the Pooling of Time Series and Cross Section Data," *Econometrica*, 46, 69–85.

Nauges, C., and A. Thomas (2003): "Consistent Estimation of Dynamic Panel Data Models with Time Varying Individual Effects," *Annales D'Economie et de Statistique*, 70, 53–75.

Nickell, S. (1981): "Biases in Dynamic Models with Fixed Effects," *Econometrica*, 49, 1417–1426.

Phillips, P., and D. Sul (2007): "Bias in Dynamic Panel Estimation with Fixed Effects, Incidental Trends and Cross-Section Dependence," *Jornal of Econometrics*, 137, 162–188.

Rothenberg, T. (1971): "Identification in Parametric Models," *Econometrica*, 39:3, 577–591.

Smith, A. (1993): "Estimating Nonlinear Time Series Models Using Simulated Vector Autoregressions," *Journal of Applied Econometrics*, 8:S1, S63–S84.

Table 1: Dynamic Panel Model:

$$y_{it} \quad = \quad \lambda_i + A_1 y_{i,t-1} + B_0 x_{it} + u_{it}, \quad \sigma^2_{u,t} = \sigma_{u0}(.95 - 0.5T + .1t), t \geq 1.$$

	true	LSDV	se	mse	BC	se	mse	IDE	se	mse
$(N,T) = (300,2)$										
β_1	1.000	0.883	0.074	0.138	1.005	0.086	0.086	0.997	0.098	0.098
α	0.800	0.353	0.049	0.449	0.837	0.071	0.080	0.802	0.060	0.060
σ_u	1.000	-	-	-	-	-	-	1.089	0.048	0.101
$(N,T) = (200,3)$										
β_1	1.000	0.967	0.064	0.072	1.004	0.067	0.067	1.012	0.081	0.082
α	0.800	0.541	0.041	0.262	0.806	0.053	0.053	0.816	0.046	0.049
σ_u	1.000	-	-	-	-	-	-	1.017	0.042	0.046
$(N,T) = (150,4)$										
β_1	1.000	0.992	0.047	0.048	1.000	0.048	0.048	1.002	0.067	0.066
α	0.800	0.635	0.028	0.168	0.810	0.032	0.034	0.823	0.033	0.041
σ_u	1.000	-	-	-	-	-	-	1.036	0.033	0.049
$(N,T) = (100,6)$										
β_1	1.000	1.013	0.046	0.048	1.000	0.046	0.046	1.002	0.062	0.062
α	0.800	0.710	0.024	0.094	0.809	0.026	0.028	0.826	0.030	0.040
σ_u	1.000	-	-	-	-	-	-	1.054	0.037	0.065
$(N,T) = (60,10)$										
β_1	1.000	1.019	0.045	0.049	0.999	0.044	0.044	1.000	0.061	0.061
α	0.800	0.759	0.020	0.046	0.809	0.020	0.022	0.810	0.024	0.026
σ_u	1.000	-	-	-	-	-	-	1.098	0.036	0.105
$(N,T) = (40,15)$										
β_1	1.000	1.017	0.041	0.044	0.996	0.041	0.041	1.001	0.058	0.058
α	0.800	0.776	0.017	0.029	0.808	0.018	0.019	0.801	0.023	0.023
σ_u	1.000	-	-	-	-	-	-	1.114	0.039	0.120

Note: BC denotes the bias-corrected in Bun and Carree (2005). IDE uses the LSDV to estimate the auxiliary parameters. LSDV denotes the least squares dummy variable (or within) estimates.

Table 2: Panel ARX(1,1): No Measurement Error ($N = 200$)

$$y_{it} \quad = \quad \lambda_i + A_1 y_{i,t-1} + B_0 x_{it} + B_1 x_{i,t-1} + u_{it}$$

	true	LSDV_y	se	mse	IDEA	se	mse	IDEA	se	mse
	$T = 5$					OLS			LSDV	
A_1	0.600	0.237	0.032	0.364	0.929	0.010	0.329	0.525	0.047	0.088
σ_u	1.000	0.149	0.007	0.851	1.153	0.010	0.153	1.032	0.055	0.063
σ_v	-	-	-	-	0.274	0.017	0.275	0.081	0.057	0.099
ϕ	-	-	-	-	0.222	0.026	0.224	0.020	0.060	0.063
B_0	1.000	1.033	0.051	0.061	1.414	0.062	0.419	0.994	0.048	0.049
A_1	0.600	0.530	0.078	0.105	0.503	0.019	0.099	0.608	0.066	0.066
σ_u	1.000	0.151	0.099	0.855	0.808	0.248	0.314	0.868	0.338	0.362
σ_{vy}	-	-	-	-	0.057	0.148	0.159	0.170	0.264	0.314
ϕ_y	-	-	-	-	0.067	0.122	0.139	0.075	0.066	0.100
B_0	1.000	0.972	0.056	0.062	0.981	0.119	0.120	1.018	0.084	0.086
B_1	0.600	0.644	0.065	0.079	0.793	0.158	0.249	0.612	0.089	0.090
A_1	0.800	0.747	0.061	0.080	0.799	0.129	0.129	0.795	0.068	0.068
σ_u	1.000	0.228	0.162	0.788	0.957	0.334	0.337	0.876	0.375	0.395
σ_v^x	-	-	-	-	0.233	0.180	0.322	0.101	0.136	0.169
ϕ_x	-	-	-	-	0.186	0.197	0.271	0.102	0.010	0.103
B_0	1.000	0.982	0.042	0.046	1.226	0.256	0.341	1.023	0.067	0.071
B_1	0.600	0.639	0.050	0.063	0.913	0.226	0.386	0.581	0.062	0.064
A_1	0.800	0.763	0.036	0.052	0.551	0.197	0.317	0.810	0.040	0.042
σ_u	1.000	0.192	0.097	0.814	0.588	0.423	0.590	0.816	0.282	0.337
σ_{vx}	-	-	-	-	0.402	0.211	0.454	0.108	0.120	0.162
ϕ_x	-	-	-	-	0.377	0.274	0.466	0.096	0.036	0.102
σ_{vy}	-	-	-	-	0.114	0.113	0.160	0.183	0.218	0.284
ϕ_y	-	-	-	-	0.105	0.044	0.114	0.085	0.082	0.118

Note: LSDV$-y$ are the infeasible estimates using the latent (true) variables. IDEA-OLS uses OLS to estimate the auxiliary model, while IDEA-LSDV uses the LSDV estimator.

Table 3a: Panel ARX(1,1) Model with Measurement Error in X_{it}, $(N,T)=(200,5)$

$$\begin{aligned} y_{it} &= \lambda_i + A_1 y_{i,t-1} + B_0 x_{it} + B_1 x_{i,t-1} + u_{it} \\ X_{it} &= x_{it} + \epsilon^x_{it}, \quad \epsilon^x_{it} = \phi_x \epsilon_{i,t-1} + v_{it}, \quad v^x_{it} \sim N(0,\sigma^2_{vx}) \\ Y_{it} &= y_{it} + \epsilon^y_{it}, \quad \epsilon^y_{it} = \phi_x \epsilon^y_{i,t-1} + v^y_{it}, \quad v^y_{it} \sim N(0,\sigma^2_{vy}). \end{aligned}$$

	true	LSDV_y	se	mse	LSDV_Y	se	mse	IDEA	se	mse
A_1	0.600	0.237	0.032	0.365	0.223	0.034	0.378	0.512	0.063	0.108
σ_u	1.000	0.149	0.007	0.851	0.187	0.009	0.813	1.040	0.050	0.064
σ_v	0.500	-	-	-	-	-	-	0.531	0.070	0.077
ϕ	0.500	-	-	-	-	-	-	0.450	0.078	0.093
B_0	1.000	1.033	0.051	0.061	1.042	0.052	0.067	0.991	0.049	0.050
A_1	0.600	0.530	0.079	0.106	0.511	0.075	0.116	0.614	0.061	0.063
σ_u	1.000	0.151	0.100	0.855	0.191	0.099	0.815	0.891	0.310	0.329
σ_{vy}	0.500	-	-	-	-	-	-	0.463	0.173	0.177
ϕ_y	0.500	-	-	-	-	-	-	0.381	0.472	0.487
B_0	1.000	0.973	0.055	0.061	0.788	0.053	0.218	1.072	0.258	0.268
B_1	0.600	0.645	0.062	0.077	0.532	0.058	0.089	0.630	0.212	0.214
A_1	0.800	0.746	0.061	0.081	0.749	0.062	0.080	0.788	0.086	0.087
σ_u	1.000	0.226	0.156	0.789	0.333	0.139	0.681	0.893	0.360	0.375
σ^x_v	0.500	-	-	-	-	-	-	0.493	0.271	0.271
ϕ_x	0.500	-	-	-	-	-	-	0.383	0.073	0.138
B_0	1.000	0.984	0.035	0.038	0.791	0.040	0.213	1.021	0.218	0.218
B_1	0.600	0.640	0.047	0.061	0.538	0.048	0.078	0.611	0.235	0.235
A_1	0.800	0.765	0.031	0.046	0.757	0.031	0.053	0.804	0.056	0.056
σ_u	1.000	0.180	0.093	0.825	0.336	0.096	0.670	1.075	0.241	0.251
σ_{vx}	0.500	-	-	-	-	-	-	0.502	0.228	0.227
ϕ_x	0.500	-	-	-	-	-	-	0.359	0.193	0.238
σ_{vy}	0.500	-	-	-	-	-	-	0.401	0.006	0.099
ϕ_y	0.500	-	-	-	-	-	-	0.401	0.003	0.099

Note: LSDV$-y$ are the infeasible estimates using the latent (true) variables. LSDV$-Y$ are estimates using the contaminated data. The IDEA estimates applies LSDV to the auxiliary model.

Table 3b: Panel ARX(1,1) Model with Measurement Error in X_{it}, $(N, T) = (200, 10)$

$$
\begin{aligned}
y_{it} &= \lambda_i + A_1 y_{i,t-1} + B_0 x_{it} + B_1 x_{i,t-1} + u_{it} \\
X_{it} &= x_{it} + \epsilon^x_{it}, \quad \epsilon^x_{it} = \phi_x \epsilon_{i,t-1} + v_{it}, \quad v^x_{it} \sim N(0, \sigma^2_{vx}) \\
Y_{it} &= y_{it} + \epsilon^y_{it}, \quad \epsilon^y_{it} = \phi_x \epsilon^y_{i,t-1} + v^y_{it}, \quad v^y_{it} \sim N(0, \sigma^2_{vy}).
\end{aligned}
$$

	true	LSDV$_y$	se	mse	LSDV$_Y$	se	mse	IDEA	se	mse
A_1	0.600	0.421	0.022	0.180	0.406	0.022	0.195	0.592	0.030	0.032
σ_u	1.000	0.088	0.003	0.912	0.110	0.004	0.890	0.924	0.036	0.084
σ_v	0.500	-	-	-	-	-	-	0.454	0.032	0.056
ϕ	0.500	-	-	-	-	-	-	0.456	0.042	0.060
B_0	1.000	1.023	0.037	0.043	1.028	0.039	0.048	0.994	0.036	0.036
A_1	0.600	0.563	0.044	0.058	0.553	0.044	0.064	0.608	0.037	0.038
σ_u	1.000	0.090	0.038	0.911	0.113	0.038	0.888	0.955	0.219	0.222
σ_{vy}	0.500	-	-	-	-	-	-	0.490	0.155	0.155
ϕ_y	0.500	-	-	-	-	-	-	0.387	0.345	0.361
B_0	1.000	0.990	0.026	0.028	0.797	0.028	0.205	0.987	0.073	0.074
B_1	0.600	0.626	0.034	0.043	0.503	0.033	0.103	0.567	0.052	0.061
A_1	0.800	0.781	0.017	0.025	0.790	0.018	0.020	0.811	0.019	0.022
σ_u	1.000	0.148	0.049	0.853	0.300	0.062	0.702	1.036	0.213	0.216
σ^x_v	0.500	-	-	-	-	-	-	0.488	0.084	0.084
ϕ_x	0.500	-	-	-	-	-	-	0.420	0.044	0.092
B_0	1.000	0.989	0.026	0.028	0.792	0.030	0.210	0.997	0.112	0.112
B_1	0.600	0.626	0.034	0.043	0.510	0.034	0.096	0.579	0.096	0.098
A_1	0.800	0.782	0.017	0.025	0.783	0.018	0.024	0.804	0.024	0.025
σ_u	1.000	0.148	0.046	0.853	0.316	0.059	0.686	1.002	0.138	0.138
σ_{vx}	0.500	-	-	-	-	-	-	0.495	0.117	0.117
ϕ_x	0.500	-	-	-	-	-	-	0.404	0.088	0.130
σ_{vy}	0.500	-	-	-	-	-	-	0.520	0.132	0.133
ϕ_y	0.500	-	-	-	-	-	-	0.400	0.121	0.157

Table 4: Panel ARX(1,0) Model with Endogenous Regressors ($N = 200$)

$$\begin{aligned} y_{it} &= \lambda_i + A_1 y_{i,t-1} + B_0 X_{it} + u_{it} \\ X_{it} &= x_{it} + \gamma u_{it}, \\ x_{it} &= \lambda_i + \rho x_{i,t-1} + u^x_{it}. \end{aligned}$$

	true	LSDV_x	se	mse	LSDV_X	se	mse	IDEA	se	mse
$T = 5,\ \rho = .5$										
B_0	1.000	0.998	0.036	0.036	0.998	0.036	0.036	1.006	0.259	0.259
A_1	0.800	0.736	0.067	0.093	0.736	0.067	0.093	0.786	0.061	0.063
σ_{u0}	1.000	-	-	-	-	-	-	0.886	0.359	0.376
γ	-	-	-	-	-	-	-	0.034	0.384	0.385
B_0	1.000	0.995	0.053	0.053	1.319	0.183	0.368	1.002	0.256	0.256
A_1	0.800	0.680	0.115	0.166	0.738	0.058	0.085	0.773	0.097	0.100
σ_{u0}	1.000	-	-	-	-	-	-	0.832	0.531	0.557
γ	0.500	-	-	-	-	-	-	0.701	0.297	0.358
$T = 10,\ \rho = .5$										
B_0	1.000	1.008	0.024	0.025	1.008	0.024	0.025	0.928	0.175	0.189
A_1	0.800	0.770	0.029	0.042	0.770	0.029	0.042	0.802	0.026	0.026
σ_{u0}	1.000	-	-	-	-	-	-	0.956	0.260	0.263
γ	-	-	-	-	-	-	-	0.114	0.243	0.269
B_0	1.000	1.016	0.036	0.040	1.333	0.131	0.358	1.003	0.130	0.130
A_1	0.800	0.742	0.054	0.080	0.761	0.030	0.049	0.794	0.029	0.030
σ_{u0}	1.000	-	-	-	-	-	-	0.934	0.309	0.315
γ	0.500	-	-	-	-	-	-	0.625	0.133	0.182

Note: LSDV$-x$ are the infeasible estimates using the orthogonal component of the regressor. LSDV$-X$ are estimates using the endogenous regressor. The IDEA estimates applies LSDV to the auxiliary model.

Table 5a: Panel Model with Cross-Section Dependence: $N = 200$ (Homoskedasticity)

$$y_{it} = \lambda_i f_t + B_{01} X_{1it} + B_{02} X_{2it} + u_{it}, \quad u_{it} \sim N(0, \sigma_u^2)$$
$$X_{kit} = x_{kit} + \lambda_i g_t + \mu_k, \quad k = 1, 2.$$

	true	OLS	se	mse	LSDV	se	mse	IDEA	se	mse
$\sigma_f^2 = 0$, $\sigma_\lambda^2 = 1$										
B_{01}	1.000	0.998	0.025	0.025	0.991	0.032	0.033	1.024	0.043	0.050
B_{02}	2.000	2.001	0.025	0.025	1.991	0.031	0.033	2.019	0.042	0.046
A_1	0.600	0.600	0.005	0.005	0.589	0.006	0.012	0.607	0.010	0.012
σ_u	1.000	-	-	-	-	-	-	0.963	0.030	0.048
$\bar{\lambda}$	0.500	-	-	-	-	-	-	0.452	0.023	0.054
σ_λ	1.000	-	-	-	-	-	-	0.921	0.026	0.084
$\sigma_f^2 = 1$, $\sigma_\lambda^2 = 1$, $T = 10$										
B_{01}	1.000	0.873	0.027	0.130	0.848	0.035	0.156	0.953	0.045	0.065
B_{02}	2.000	1.843	0.026	0.159	1.848	0.033	0.156	1.947	0.044	0.069
A_1	0.600	0.643	0.005	0.043	0.644	0.007	0.045	0.610	0.010	0.014
σ_u	1.000	-	-	-	-	-	-	1.000	0.032	0.032
$\bar{\lambda}$	0.500	-	-	-	-	-	-	0.471	0.063	0.070
σ_λ	1.000	-	-	-	-	-	-	0.947	0.063	0.082
$\sigma_f^2 = 1$, $\sigma_\lambda^2 = 2$, $T = 5$										
B_{01}	1.000	0.820	0.028	0.182	0.786	0.039	0.217	0.946	0.044	0.070
B_{02}	2.000	1.776	0.027	0.226	1.783	0.036	0.220	1.935	0.043	0.078
A_1	0.600	0.660	0.005	0.060	0.665	0.007	0.065	0.613	0.010	0.016
σ_u	1.000	-	-	-	-	-	-	1.017	0.035	0.039
$\bar{\lambda}$	0.500	-	-	-	-	-	-	0.484	0.053	0.055
σ_λ	1.000	-	-	-	-	-	-	0.969	0.047	0.056
$\sigma_f^2 = 1$, $\sigma_\lambda^2 = 1$, $T = 10$										
B_{01}	1.000	0.980	0.020	0.028	0.977	0.024	0.033	0.951	0.029	0.057
B_{02}	2.000	1.950	0.019	0.053	1.964	0.023	0.042	1.952	0.030	0.057
A_1	0.600	0.638	0.004	0.038	0.642	0.004	0.042	0.590	0.007	0.012
σ_u	1.000	-	-	-	-	-	-	0.971	0.017	0.033
$\bar{\lambda}$	0.500	-	-	-	-	-	-	0.575	0.051	0.090
σ_λ	1.000	-	-	-	-	-	-	0.822	0.071	0.191
$\sigma_f^2 = 1$, $\sigma_\lambda^2 = 2$,$T = 10$										
B_{01}	1.000	0.970	0.023	0.038	0.966	0.027	0.043	0.938	0.032	0.070
B_{02}	2.000	1.926	0.021	0.077	1.946	0.026	0.060	1.937	0.032	0.071
A_1	0.600	0.655	0.004	0.055	0.662	0.005	0.062	0.592	0.007	0.011
σ_u	1.000	-	-	-	-	-	-	0.979	0.017	0.027
$\bar{\lambda}$	0.500	-	-	-	-	-	-	0.621	0.053	0.132
σ_λ	1.000	-	-	-	-	-	-	0.817	0.057	0.192

Note: OLS denotes least squares estimates of the panel model. LSDV denotes the within estimates of the model. The IDEA estimates applies OLS to the auxiliary model.

Table 5b: Panel Model with Cross-Section Dependence: $N = 200$ (Heteroskedasticity)

$$y_{it} = \lambda_i f_t + B_{01} X_{1it} + B_{02} X_{2it} + u_{it}, \quad u_{it} \sim N(0, \sigma_{ut}^2)$$
$$X_{kit} = x_{kit} + \lambda_i g_t + \mu_k, \quad k = 1, 2.$$

	true	OLS	se	mse	LSDV	se	mse	IDEA	se	mse
$\sigma_f^2 = 0$, $\sigma_\lambda^2 = 1$	$\sigma_{ut} = .95 - .05T + .1t$, $T = 5$									
B_{01}	1.000	0.999	0.023	0.023	0.995	0.039	0.039	1.022	0.037	0.043
B_{02}	2.000	2.001	0.025	0.025	1.996	0.037	0.037	2.023	0.037	0.044
A_1	0.600	0.600	0.005	0.005	0.588	0.007	0.014	0.601	0.011	0.011
σ_u	1.000	-	-	-	-	-	-	1.012	0.035	0.037
$\bar{\lambda}$	0.500	-	-	-	-	-	-	0.447	0.017	0.055
σ_λ	1.000	-	-	-	-	-	-	0.953	0.034	0.058
$\sigma_f^2 = 1$, $\sigma_\lambda^2 = 1$	$\sigma_{ut} = .95 - .05T + .1t$, $T = 5$									
B_{01}	1.000	0.890	0.024	0.113	0.845	0.042	0.161	0.976	0.038	0.045
B_{02}	2.000	1.833	0.026	0.169	1.831	0.040	0.174	1.956	0.042	0.061
A_1	0.600	0.647	0.005	0.047	0.647	0.007	0.048	0.616	0.010	0.019
σ_u	1.000	-	-	-	-	-	-	1.093	0.033	0.098
$\bar{\lambda}$	0.500	-	-	-	-	-	-	0.494	0.096	0.097
σ_λ	1.000	-	-	-	-	-	-	0.933	0.115	0.133
$\sigma_f^2 = 1$, $\sigma_\lambda^2 = 2$	$\sigma_{ut} = .95 - .05T + .1t$, $T = 5$									
B_{01}	1.000	0.845	0.025	0.157	0.779	0.045	0.226	0.981	0.039	0.044
B_{02}	2.000	1.766	0.028	0.236	1.759	0.043	0.245	1.959	0.043	0.059
A_1	0.600	0.665	0.005	0.065	0.669	0.008	0.069	0.618	0.010	0.020
σ_u	1.000	-	-	-	-	-	-	1.108	0.037	0.115
$\bar{\lambda}$	0.500	-	-	-	-	-	-	0.544	0.089	0.099
σ_λ	1.000	-	-	-	-	-	-	1.019	0.070	0.073
$\sigma_f^2 = 0$, $\sigma_\lambda^2 = 1$	$\sigma_{ut} = .95 - .05T + .1t$, $T = 10$									
B_{01}	1.000	0.959	0.018	0.044	0.968	0.027	0.041	0.961	0.039	0.055
B_{02}	2.000	1.908	0.018	0.093	1.936	0.025	0.068	1.911	0.043	0.099
A_1	0.600	0.636	0.004	0.036	0.644	0.005	0.044	0.625	0.034	0.042
σ_u	1.000	-	-	-	-	-	-	1.153	0.080	0.173
$\bar{\lambda}$	0.500	-	-	-	-	-	-	0.531	0.082	0.087
σ_λ	1.000	-	-	-	-	-	-	0.880	0.082	0.145
$\sigma_f^2 = 1$, $\sigma_\lambda^2 = 2$	$\sigma_{ut} = .95 - .05T + .1t$, $T = 10$									
B_{01}	1.000	0.940	0.019	0.064	0.951	0.030	0.057	0.922	0.027	0.083
B_{02}	2.000	1.866	0.020	0.135	1.906	0.028	0.098	1.910	0.028	0.094
A_1	0.600	0.652	0.004	0.053	0.666	0.005	0.066	0.607	0.006	0.009
σ_u	1.000	-	-	-	-	-	-	1.105	0.022	0.107
$\bar{\lambda}$	0.500	-	-	-	-	-	-	0.584	0.044	0.095
σ_λ	1.000	-	-	-	-	-	-	0.947	0.049	0.072