Confidence Intervals in Autoregressive Panels, Invariance, and Indirect Inference

Lynda Khalaf* Charles J. Saunders†‡

Carleton University Carleton University

December 22, 2014

Abstract

Confidence regions are proposed for autoregressive panels, with valid coverage whether autoregressive roots are at or close to the unit boundary, or far from unity. Four strategies are exploited: a discontinuous definition of initial observations ensures location-scale invariance; indirect inference is modified allowing for confidence set estimation; the Monte Carlo test principle supplies finite sample tests, as in contrast to bootstrapping, the existence of a limiting distribution is not required; and the objective function is inverted rather than minimized. Our combination of inversion and provable invariance of the indirect inference objective function ensures level control and good power in simulations.

JEL Classification: C13, C15, C23, C33

Keywords: Dynamic Panel Model, Indirect Inference, Monte Carlo Tests, Test Inversion, Unit Roots

*Economics Department and Centre for Monetary and Financial Economics (CMFE), Carleton University; Centre interuniversitaire de recherche en économie quantitative (CIREQ), and Groupe de recherche en économie de l'énergie, de l'environnement et des ressources naturelles (GREEN), Université Laval. Mailing address: Economics Department, Carleton University, Loeb Building 1125 Colonel By Drive, Ottawa, Ontario, K1S 5B6 Canada. Tel: (613) 520-2600-8697; email: Lynda_Khalaf@carleton.ca.

†Economics Department, Carleton University. Mailing address: Economics Department, Carleton University, Loeb Building 1125 Colonel By Drive, Ottawa, Ontario, K1S 5B6 Canada. email: charles.saunders@carleton.ca.

‡The authors would like to acknowledge and thank, Russell Davidson, Antoine Alex Djogbenou and Alexandru Minea for comments and discussion at and leading up to the 2014 CEA Conference. We also appreciate the contribution of Maurice Bun for drawing our attention to a paper on invariance and the LSDV estimator.

1 Introduction

The least squares dummy variable (LSDV) estimate of an autoregressive panel is inconsistent, for a given number of time periods. The source of the bias is an incidental parameter problem, first identified by Neyman and Scott (1948), since the lagged dependent variable is correlated with the time-invariant component of the model error. Demeaning and first differences alone are ineffective at eliminating this bias.

A common method of bias correction is based in the instrumental variable (IV) and generalized method of moment (GMM) literature, the most popular of these estimators were introduced by Anderson and Hsiao (1982), Arellano and Bond (1991), Arellano and Bover (1995), Ahn and Schmidt (1995), and Blundell and Bond (1998). The instrumental variables used are exogenous variables (if present) and estimator specific combinations of level and first difference lags of the dependent variable, and lags of the error estimates. These estimators generally rely on Large-N and Small-T for both consistency and good asymptotic properties. As shown in Bun and Kleibergen (2013) and the references therein, regularity conditions may hold only weakly, especially with Small-T settings, which may cause non-uniform convergence problems even when the number of cross-sections is large. Weak-instruments concerns complicate limiting theory, a problem not restricted of course to the panel IV case.

Another method of obtaining consistent estimates is a well defined bias-correction term, that is a function of the LSDV estimate from the data and the number of time periods. The bias-correction term is derived under either Large-N and Fixed-T conditions, in the case of Kiviet (1995) and Bun and Carree (2005), or in a Large-N and Large-T setting, for Hahn and Kuersteiner (2002) and Hahn and Moon (2006). In Monte Carlo experiments, these estimators have been shown to outperform the GMM/IV alternatives but not for all simulation settings.

A third method is that of Gouriéroux, Phillips, and Yu (2010), extending the indirect inference estimator (IIE) to the LSDV estimate of a panel-AR(1) model. IIE is a simulation-based method that generates a set of dynamic panel data under the null model and for a given set of parameters. The intuition of the IIE is that the bias term of the original data is a function of the true lagged dependent parameter and the number of time periods (T), as shown in Nickell (1981). Simulating the bias requires an assumption about the true value of the parameter and some distributional assumptions regarding the errors. The difference between the data bias and the simulated bias can be minimized, resulting in a consistent and unbiased estimate of the true parameter, under the null.

The dynamic panel literature has focused primarily on consistent estimators. In most cases, these estimators impose stationarity to ensure appropriate asymptotic properties. Unit root panel models have, for the most part, followed a separate research thread and crossing the stationary thread on occasion. Even more generally, outside the panel context, a method that works at or close to the boundary may fail elsewhere. Phillips (2014) shows that constructing confidence intervals by inverting an alternative model can result in not only poor coverage properties but even complete lack of coverage. The asymptotic properties of stationary, local-to-unity, and unit root models differ greatly, which is a serious impediment to developing a unified asymptotic theory.

Motivated by this our paper makes three contributions. Our first contribution is the introduction of a discontinuous panel-AR(1) framework suitable for joint consideration of stationary, local-to-unity, and unit root panels. The initial observation of the series interferes with the LSDV estimate of the lagged dependent parameter, hence it is a nuisance parameter. We demonstrate that ignoring the nuisance parameter causes serious size distortions. To solve this problem, we propose a discontinuous functional specification of the initial observation in terms of the lagged dependent parameter as well as other hyper-parameters, specifically unobserved heterogeneity and the variance of the time-varying error. Our initial observation specification respects current practice in the literature and leads to an objective function that is provably location-scale invariant with analytical proofs provided. Our proposed approach demonstrates that a discontinuity that is an impediment to unifying asymptotic theory can be an asset in a finite-sample context.

Invariance proofs in the dynamic panel context are limited. Kiviet and van den Doel (1994) prove invariance of a LSDV estimator that has been augmented with an additional lag of exogenous regressors. Their augmented-LSDV estimator relies on exogenous regressors for invariance and four or more time periods.

Our second contribution extends the indirect inference estimator in Gouriéroux, Phillips, and Yu (2010) to jointly consider stationary and unit root panels, through the inclusion of our discontinuous framework. Standard objective functions are ill-behaved at or close to unity because the long-run variance is not finite. Our modification corrects this problem without costing power. In contrast to point estimation through minimization, we focus on confidence set estimation by inverting the objective function.

Formally, our proposed method treats the indirect inference objective function as the test statistic which we invert rather than minimize. Inversion of our test statistic involves collecting the lagged dependent parameter values that are not rejected at a given level. We generate two sets of simulated series under our joint null, which is comprised of a given lagged dependent parameter value and our discontinuous initial observation function. The first set of simulated series are used to define out test statistic that is computed as the objective function realization from the data under the null. The test statistic is standardized by the second set of simulated series and a simulation-based p-value is obtained by Monte Carlo methods [refer to Dufour (2006), Dufour and Valéry (2009) and Beaulieu, Dufour, and Khalaf (2007, 2013, 2014)]. We show that the Monte Carlo technique allows for discontinuity under the null and works to construct exact confidence sets, even though conventional methods could lead to highly inaccurate inference. Our proposed method is an interesting contribution to the theory of test inversion and simulation-based methods.

Our third contribution examines the properties of the indirect confidence set inference method through a simulation study. The indirect confidence set inference method exhibits level control and power, even under modest cross-sections and small number of time periods conditions. The level control and power properties of the method are retained for stationary and unit root autoregressive panels. For the stationary panel-AR(1) simulations, as a means of comparison we examine the coverage of the bootstrap method of Gonçalves and Kaffo (2014).

The paper is structured as follows, section 2 details our panel-AR(1) model, assumptions, parameter estimation, invariance proofs and bias. Section 4 presents the indirect inference method to the panel-AR(1) model, and extends it to allow for the assumption of discontinuous initial value function. Also in this section, we introduce the indirect confidence set inference method to the panel-AR(1) model to obtain an exact confidence set for the parameter. In section 5, we present the design and results of our simulation study for our method and find level control and good power properties. Section 6 summarizes the major findings and indicates some potential extensions to non-Gaussian disturbances.

2 Framework

The general form of a panel-AR(1) model is given as,

$$\begin{aligned} y_{i,t} &= (1-\phi)\,\mu_i + \phi y_{i,t-1} + \epsilon_{i,t} \\ & i = 1,\ldots,N \\ & t = 1,\ldots,T, \end{aligned} \tag{1}$$

where i indicates the cross-section of the panel, t is the time index of the observations, and $y_{i,t}$ is the dependent variable. Alternative representations of this model have been implemented in the literature and are discussed below. The representation given above is robust to roots of the lag polynomial that lie outside or on the unit circle. The $\epsilon_{i,t}$ are time-varying errors, which follow a distribution centered at zero that allows for independent draws. The time-invariant μ_i terms can either be non-random or random disturbances with the first two moments defined as μ and σ^2_μ. The existence of these moments does not intervene in our proposed approach presented below since our theorems do not rely on asymptotics but instead a finite-sample approach. It is convention that $\epsilon_{i,t}$ and μ_i are contemporaneously uncorrelated,

$$E[\epsilon_{i,t}|\mu_j] = 0 \;\; \forall\; i,j.$$

Common in the literature is the assumption that both errors follow a Gaussian distribution, but clearly the framework is robust to most alternative distributional assumptions.

2.1 LSDV and Bias

The panel-AR(1) model presented in equation 1 in demeaned form is

$$(y_{i,t} - y_{i,\bullet}) = \phi(y_{i,t-1} - y_{i,\bullet-1}) + (\epsilon_{i,t} - \epsilon_{i,\bullet})$$

where the sample means are given by

$$y_{i,\bullet} = \frac{1}{T-1}\sum_{t=2}^{T} y_{i,t}, \quad y_{i,\bullet-1} = \tfrac{1}{T-1}\sum_{t=1}^{T-1} y_{i,t}, \text{ and } \quad \epsilon_{i,\bullet} = \frac{1}{T-1}\sum_{t=2}^{T}\epsilon_{i,t}.$$

A condition of the LSDV estimator is that minimum number of time periods is defined by $T \geq 3$. We set a distinction here, so for the models presented the dependent variable is observed for $t \geq 1$, and not observed for $t \leq 0$, so LSDV estimate is based on the observed series.[1] To simplify the notation, the demeaned series will be denoted with an asterisk $(*)$. So the estimated equation is

$$y^*_{i,t} = y^*_{i,t-1}\phi + \epsilon^*_{i,t}$$

and through appropriate stacking we obtain the matrix form,

$$y^* = y^*_{-1}\phi + \epsilon^*. \tag{2}$$

The LSDV estimates of ϕ for equation 2 gives

$$\hat{\phi} = \left(y^{*\prime}_{-1}y^*_{-1}\right)^{-1} y^{*\prime}_{-1}y^* = \phi + \left(y^{*\prime}_{-1}y^*_{-1}\right)^{-1} y^{*\prime}_{-1}\epsilon^*. \tag{3}$$

The estimate of ϕ is biased due to $E[y^{*\prime}_{-1}\epsilon^*] \neq 0$ for fixed-T, and this bias is analytically derived in Nickell (1981).

2.2 Initial Observation

The unobserved lead values affect parameter estimation in dynamic models since their influence persists. The general assumption is that the series was initiated long enough in the distant past so that the effects of the initial observation are negligible.

A historical approach was to simulate a series that was longer than required then retain only the most recent observations, an example of this approach is found in the Monte Carlo study of Arellano and Bond (1991). However, the number of truncated observations depends on the

[1] An alternative approach divides the observed series into modeled and initial observations with notionally equivalent implications.

model parameters, as parameter settings approach nonstationarity the number of required truncated observations increases.

Alternatively, the initial observation can be specified parametrically, which relaxes the need to generate and truncate additional lead observations. Our parametric specification is given by assumption 1 and is a general in the sense that it encompasses the majority of parametric specifications used in the literature. The dispersion of the initial observation is determined by σ_γ.

Assumption 1 (Initial Values for Simulated Series) *The initial value of the series for any i is defined by*

$$y_{i,0} = \mu_i + \sigma_\gamma \epsilon_{i,0}$$

where $\sigma_\gamma \geq 0$, and $\epsilon_{i,0}$ follows the same distribution as $\epsilon_{i,t}$ and $E[\epsilon_{i,0}\epsilon_{i,t}] = 0$ for all t.

In finite sample and standard time series approaches it is common to condition on the initial observation by setting the dispersion of the initial observation to zero.

Assumption 2 (Conditional σ_γ) *Let the initial value of the series for any i under assumption 1 be a function of $\sigma_\gamma(\phi)$, which is defined as*

$$\sigma_\gamma := \sigma_\gamma(\phi) = 0 \text{ for all } \phi.$$

Assumption 2 means that the initial observation is equal to the unobserved heterogeneity of the series, $y_{i,0}(\phi) = \mu_i$. In the dynamic panel context, Gonçalves and Kaffo (2014) impose an initial value setting under no dispersion,

$$\begin{aligned} y_{i,t} &= a_i + \phi y_{i,t-1} + \epsilon_{i,t} \text{ and} \\ y_{i,0} &= \frac{a_i}{1-\phi}. \end{aligned}$$

This setting can be transformed into equation 1 by substituting the following identity, $a_i = (1-\phi)\mu_i$.

The importance of assumptions on the initial observation were numerically examined in a GMM framework by Hahn (1999) finding strong evidence in favour of imposing stationarity. To account for the unit root, we introduce a discontinuous null setting for σ_γ that is a function of ϕ.

Assumption 3 (Discontinuous σ_γ) *Let the initial value of the series for any i under assumption 1 be a function of $\sigma_\gamma(\phi)$, which is defined as*

$$\sigma_\gamma := \sigma_\gamma(\phi) = \begin{cases} \frac{1}{\sqrt{(1-\phi^2)}} & \textit{if } |\phi| < 1, \\ 0 & \textit{if } \phi = 1. \end{cases}$$

Under assumption 3, the initial observation is discontinuous in ϕ, specifically

$$\begin{aligned} y_{i,0}(\phi) &= \mu_i + \sigma_\gamma(\phi)\epsilon_{i,0} \\ &= \begin{cases} \mu_i + \frac{1}{\sqrt{(1-\phi^2)}}\epsilon_{i,0} & \text{if } |\phi| < 1, \\ \mu_i & \text{if } \phi = 1. \end{cases} \end{aligned}$$

As $\phi \longrightarrow 1$ then the initial observation is ill-defined, $y_{i,0}(\phi) \longrightarrow \infty$, which affects the parameter estimates unless a functional form can account for the discontinuity at unity. These problems at the limit are avoided under assumption 3. Under assumption 3, at unity the initial observation is only a function of the time-invariant error, corollary 2 justifies our approach since the LSDV estimate of ϕ for a unit root series is invariant to σ_γ. Under assumption 3, σ_γ is no longer a nuisance parameter since it is completely defined under the null.

Assumption 1 is a general specification that encompasses available initial value settings, below we provide a brief survey of the data generation processes available in the literature. We show that these can be defined as special cases of assumption 3, for stationary dynamic panels. Hahn and Moon (2006) use the following settings,

$$\begin{aligned} y_{i,t} &= (1-\phi)\mu_i + \phi y_{i,t-1} + \epsilon_{i,t} \text{ and} \\ y_{i,0} &\sim \mathcal{N}\left(\mu_i, \frac{1}{1-\phi^2}\right). \end{aligned}$$

Setting $\epsilon_{i,0} \sim \mathcal{N}(0,1)$ demonstrates this as a special case of assumption 1.

Gouriéroux, Phillips, and Yu (2010), Bun and Carree (2005), and Hahn and Kuersteiner (2002) assume the following set of equations,

$$\begin{aligned} y_{i,t} &= a_i + \phi y_{i,t-1} + \epsilon_{i,t} \text{ and} \\ y_{i,0} &= \frac{a_i}{1-\phi} + \frac{\epsilon_{i,0}}{\sqrt{1-\phi^2}}, \end{aligned}$$

where the initial observation is defined by

$$y_{i,0} \sim \mathcal{N}\left(\frac{a_i}{1-\phi}, \frac{\sigma_\epsilon^2}{1-\phi^2}\right).$$

This framework can be transformed into our framework, $a_i = (1-\phi)\mu_i$, and the initial value is obtained by assuming $\epsilon_{i,0} \sim \mathcal{N}\left(0, \sigma_\epsilon^2\right)$, and $\sigma_\gamma = \left(1-\phi^2\right)^{-1/2}$. Kiviet and van den Doel (1994) impose this special case when no exogenous regressors are present.

The papers cited above, post-date the analytical bias derived in Nickell (1981). Table 1 shows that the Nickell bias of the LSDV estimator is sensitive to assumptions on the magnitude of σ_γ.

Table 1: LSDV Bias Sensitivity to σ_γ

	σ_γ					
ϕ	0	1	1.25	1.5	2	3
0	-0.2148	-0.2148	-0.2148	-0.2148	-0.2148	-0.2148
0.5	-0.4496	-0.4346	-0.4269	-0.4179	-0.3977	-0.3544
0.6	-0.5086	-0.4901	-0.4807	-0.4699	-0.4461	-0.3971
0.7	-0.5694	-0.5506	-0.5410	-0.5300	-0.5056	-0.4553
1	-0.7153	-0.7153	-0.7153	-0.7153	-0.7153	-0.7153

Based on 1000 Monte Carlo replications, $N = 200$, $T = 5$,
and standard deviation of the bias ranges from 0.0006 to 0.0009.

For $\phi = 0$ and $\phi = 1$ the value of σ_γ does not change the value of the LSDV estimates. For $0 < \phi < 1$, as the value of σ_γ increases the bias is reduced, so σ_γ is a nuisance parameter for the LSDV bias. The table above shows that in the local region of a given ϕ the bias is not unique when σ_γ is a free parameter. The initial observation can lead to large distortions in the LSDV estimator properties without a coherent specification of σ_γ. In section 5, we revisit the σ_γ nuisance parameter problem in our simulation study and find size distortions and indications that power also suffers.

3 Invariance

The recursion of the series is the building block for the proofs and corollaries, so it is stated here as Lemma 1.

Lemma 1 (Recursive Decomposition) *The demeaned series by recursive substitution are*

$$\begin{aligned} y^*_{i,t} &= y_{i,t} - y_{i,\bullet} = g_i(\phi, \sigma_\gamma, \epsilon_{i,0}, \epsilon_i, t, T), \text{ where} \\ g_i(\phi, \sigma_\gamma, \epsilon_{i,0}, \epsilon_i, t, T) &= \left(\sum_{r=0}^{t-1} \epsilon_{i,t-r}\phi^r - \frac{1}{T-1} \sum_{s=2}^{T} \sum_{r=0}^{s-1} \epsilon_{i,t-r}\phi^r \right) \\ &\quad + \left(\sigma_\gamma \epsilon_{i,0}\phi^t - \frac{1}{T-1} \sum_{s=2}^{T} \sigma_\gamma \epsilon_{i,0}\phi^s \right). \end{aligned} \tag{4}$$

Let $g_{i,t} = g_i(\phi, \sigma_\gamma, \epsilon_{i,0}, \epsilon_i, t, T)$ *and* $g_{i,t-1} = g_i(\phi, \sigma_\gamma, \epsilon_{i,0}, \epsilon_i, t-1, T)$*, and the vector representations are* g *and* g_{-1}*, respectively. Then, the LSDV estimate of* ϕ *is*

$$\tilde{\phi} = (g'_{-1}g_{-1})^{-1}g'_{-1}g = \phi + (g'_{-1}g_{-1})^{-1}g'_{-1}\epsilon^*. \tag{5}$$

Under assumption 1 and 3 and lemma 1, the parameter estimate from equation 3 is numerically equal to the estimate from equation 5, or $\tilde{\phi} = \hat{\phi}$. This implies that the distribution of $\hat{\phi}$ is completely determined by current, lagged and lead errors (ϵ), the true value of ϕ, σ_γ, and unobserved random disturbances ($\epsilon_{i,0}$).

In the context of this paper, an estimator is invariant when a known data transformation results in an unchanged parameter estimate. A scale transformation is when the data multiplied by a positive constant, and a location transformation is the addition of a non-zero constant to either the entire series or individual cross-sections. In the LSDV framework, we focus on invariance in terms of the lagged dependent parameter (ϕ) estimate.

Theorem 1 (Scale and Location Invariance) *Under assumption 1 and for a given* σ_γ *under assumption 3, the least-square dummy variable estimate of* ϕ *from a model defined by*

$$y_{i,t} = (1-\phi)\mu_i + \phi y_{i,t-1} + \epsilon_{i,t},$$

is scale invariant in terms of σ_ϵ and location invariant in terms of μ_i.

Proof of Theorem 1 is provided in the appendix.

The LSDV estimate of ϕ is a function of the true parameter value, σ_γ, and a set of hyper-parameters, specifically: σ_ϵ^2, σ_μ^2, and μ. By theorem 1, the LSDV estimator is invariant to the hyper-parameters. By assumption 3, we define σ_γ as a discontinuous function of lagged dependent parameter. Although seemingly trivial, in our approach the initial observation is completely defined by the lagged dependent parameter and the hyper-parameters and as such the estimate of ϕ does not change. In the context of our simulation study presented in section 5, we examine numerically the implications of fixing σ_γ to a different non-negative value and violating assumption 3.

Apparent by Lemma 1, the estimate of ϕ depends on σ_γ. As $\sigma_\gamma \longrightarrow \infty$, the demeaned series is dominated by the initial observation so the LSDV estimate of ϕ will tend to the true value as the relative importance of the bias is reduced.

Corollary 1 *Under assumptions 1 and 3, the LSDV estimate is invariant to affine transformations of the data.*

Corollary 1 states that $\tilde{\phi} = \hat{\phi}$, where $\tilde{\phi}$ is the LSDV estimate from affine transformed data ($\tilde{y}_{i,t} = \kappa y_{i,t} + \delta_i$ and $\kappa \neq 0$), and $\hat{\phi}$ is the LSDV estimate from from the untransformed data ($y_{i,t}$).

Corollary 2 *Under assumptions 1, the LSDV estimate is invariant to σ_γ for $\phi = 1$.*

The implication of corollary 2 is that the LSDV estimate, $\hat{\phi}$, is not a function of σ_γ, so unchanged for any value of σ_γ as shown in table 1. Corollaries 1 and 2 are exploited from a test inversion perspective in our proposed method presented below.

4 Discontinuity, Indirect Inference, and Confidence Set

The indirect inference estimator (IIE) is a simulation-based method introduced by Gourieroux, Monfort, and Renault (1993) and independently by Smith (1993) and has been primarily used for bias correction. Gouriéroux, Phillips, and Yu (2010) apply the IIE to the panel-AR(1) context and show that their approach provides asymptotically consistent estimates of ϕ.

The LSDV estimate of ϕ is obtained from the data and denoted as $\hat{\phi}_o$. We propose the following modification by setting a joint null,

$$H_s : \phi = \phi_s \text{ and } \sigma_\gamma = \sigma_\gamma(\phi_s),$$

to a reasonable range of parameter values, $\phi_s \in \Phi = (-1, 1]$. This joint null does not lack generality, since it covers the specific settings defined in the literature as special cases. Under the joint null, we generate H simulated series,

$$\begin{aligned} y_{i,0,h} &= \sigma_\gamma(\phi_s)\epsilon_{i,0,h} \\ y_{i,t,h} &= \phi_s y_{i,t-1,h} + \epsilon_{i,t,h}. \end{aligned}$$

where $h = 1, \ldots, H$, and $\epsilon_{i,0,h}$ and $\epsilon_{i,t,h}$ are independent draws from a given distribution. We set $\mu_{i,h} = 0$ for all i, which is suitable by our invariance theorem.

For each of simulated series we compute the associated LSDV parameter estimate, denoted $\hat{\phi}_{s,h}$:

$$\hat{\phi}_{s,h} = \left(y^{*\prime}_{-1,h} y^*_{-1,h}\right)^{-1} y^{*\prime}_{-1,h} y^*_h.$$

Each simulated parameter estimate is conditional on the value of ϕ_s used in the simulation DGP and the draws of the time-varying disturbances. Under our joint null, as shown above, $\hat{\phi}_{s,h}$ are location-scale invariant for all h, hence nuisance parameter free.

The indirect inference objective function is a distance measure, represented by $\|\cdot\|$, between the estimate of ϕ from the data and a binding function based on the estimates from the simulated series, specifically

$$Q(\phi_s) = \left\| \hat{\phi}_o - b_T(\hat{\Phi}_{s,H}) \right\|.$$

Simulating under the null, relaxes the requirements that the binding function is analytically tractable, invertible, or continuous. Defining the binding function as

$$b_T(\hat{\Phi}_{s,H}) = E[\hat{\phi}_{s,h}],$$

allows for consistent estimation as $H \to \infty$ by the empirical average from the simulated series,

$$b_T(\hat{\Phi}_{s,H}) = \frac{1}{H} \sum_{h=1}^{H} \hat{\phi}_{s,h}.$$

The resulting indirect inference objective function is

$$Q(\phi_s) = \left\| \hat{\phi}_o - \frac{1}{H} \sum_{h=1}^{H} \hat{\phi}_{s,h} \right\|.$$

Gouriéroux, Phillips, and Yu (2010) applied a distance metric in a sandwich form with the identity matrix as the weighting matrix:[2]

$$Q(\phi_s) = \left(\hat{\phi}_o - \frac{1}{H} \sum_{h=1}^{H} \hat{\phi}_{s,h} \right)' \left(\hat{\phi}_o - \frac{1}{H} \sum_{h=1}^{H} \hat{\phi}_{s,h} \right). \tag{6}$$

Theorem 2 (Objective Function Location-Scale Invariant) *Under assumptions 1 and 3, and the joint null, defined as $H_s : \phi = \phi_s$ and $\sigma_\gamma = \sigma_\gamma(\phi_s)$, the distribution of $Q(\phi_s)$ is completely determined by ϕ_s and all the disturbances.*

Theorem 2 means that the indirect inference objective function is fully defined by the parameter estimates from the data and the simulated series. Indeed by theorem 1, the parameter estimates and the indirect inference objective function are nuisance parameter free, hence the objective function is pivotal under the joint null.

The information content of the binding function on the true parameter affects the quality of the estimate. Weakly defined binding functions result in less efficient bounds on the confidence set, which can occur when nuisance parameters are present, or whether the binding function is locally or globally injective. By theorem 2, the indirect inference objective function is globally injective under the joint null.

We next proceed from a confidence set perspective. Formally, we build the exact confidence set for ϕ, which is valid even for our discontinuous objective functions. Our approach inverts the indirect inference objective function, equation 6, to compute the p-value of ϕ at a given parameter

[2] Alternatives include: the Euclidean norm, and the absolute value of the distance. The absolute value may have desirable properties if the distribution of $\hat{\phi}_{s,h}$ is heavily skewed (asymmetric).

value. For a given value of ϕ, say $\phi_s \in (-1, 1]$, we compute the p-value based on the Monte Carlo methodology outlined in Dufour (2006). We build the confidence set by retaining the given values of ϕ with a p-value greater than our desired significance level. For our approach, the distribution of indirect inference objective function under the null need not be known, but only that it can be simulated.

Under the joint null, we generate M Monte Carlo replicated series and respective LSDV estimate,

$$\begin{aligned} y_{i,0,m|s} &= \sigma_\gamma(\phi_s)\epsilon_{i,0,m} \\ y_{i,t,m|s} &= \phi_s y_{i,t-1,m|s} + \epsilon_{i,t,m} \\ \hat{\phi}_{s,m} &= \left(y^{*\prime}_{-1,m} y^{*}_{-1,m}\right)^{-1} y^{*\prime}_{-1,m} y^{*}_{m}, \end{aligned}$$

where $\epsilon_{i,0,m}$ and $\epsilon_{i,t,m}$ are drawn from the same distribution as $\epsilon_{i,0,h}$ and $\epsilon_{i,t,h}$, and $\mu_{i,m} = 0$.

The objective function for each Monte Carlo series is

$$Q(\phi_{m|s}) = \left(\hat{\phi}_{s,m} - \frac{1}{H}\sum_{h=1}^{H}\hat{\phi}_{h|s}\right)' \left(\hat{\phi}_{s,m} - \frac{1}{H}\sum_{h=1}^{H}\hat{\phi}_{h|s}\right),$$

which are collected to build a set of M Monte Carlo objective function realizations, denoted as $Q_M(\phi_s)$. The set $Q_M(\phi_s)$ is used to calibrate the objective function realization from the data under the null.

Under the null, the parameter settings and assumptions on the disturbances are identical for the Monte Carlo replications and simulated series. All parameter estimates from these series will be independent and identically distributed, hence any permutation of the combined set can be used to construct $\frac{1}{H}\sum_{h=1}^{H}\hat{\phi}_{h|s}$ and each $Q(\phi_{m|s})$.

To compute the p-value under the null, we extend the method of Dufour (2006) to our statistic. We begins by counting the number of elements in $Q_M(\phi_s)$ that exceed $Q(\phi_{o|s})$, via a boolean check,

$$\text{count}(Q(\phi_{o|s})) = \sum_{m=1}^{M} \mathbf{1}\left(Q(\phi_{o|s}) \geq Q(\phi_{m|s})\right)$$

where $\mathbf{1}(\cdot)$ is one when the statement is true and zero otherwise. The p-value is computed as

$$\text{p-value}(Q(\phi_{o|s})) = \frac{1 + \text{count}(Q(\phi_{o|s}))}{1 + M}.$$

The confidence interval for ϕ at significance level α is constructed by collecting all the values of ϕ that satisfy the right-tailed condition,

$$CI(\phi|\alpha) = \left\{\phi_s \left| \text{p-value}(Q(\phi_{o|s})) > \alpha \right.\right\}.$$

The estimate of ϕ from the data and simulations are free of nuisance parameters, and conditional on a given N and T. By theorem 2, the indirect inference objective functions are pivotal so the confidence interval will be exact. Since the simulated series and Monte Carlo replications are exchangeable under the joint null, then the confidence interval will have level control, in the sense of

$$Pr(\phi \in CI(\phi|\alpha)) = 1 - \alpha.$$

Alternative formulations of IIE have been suggested, that include score-based by Gallant and Tauchen (1996), likelihood-distance-based by Smith (1993), and simulated GMM. Each of the three IIE methods are optimal under certain conditions, although the use of the *sample mean* is the least restrictive but most computationally intensive.

5 Simulation Study

We use a simulation study to examine the rejection frequency properties of the indirect confidence set inference method. Our method will be size correct under the null by construction, and the rejection frequency under the alternative (power) is of interest for ϕ coverage.

The LSDV estimations are invariant to both location and scale, under assumptions 1 and 3, so our simulation design needs specify ϕ, and the distribution of ϵ used for draws for the H simulated and M Monte Carlo simulated data.

For our tests we set parameter values that are non-dynamic, stationary, local-to-unity, and unit root, $\phi = \{0.0, 0.6, 0.9, 0.999, 1\}$. We build the confidence set under the joint null, outlined in the previous section, where $\phi_g \in (-1, 1]$. As a simulation study, we can limit our rejection frequencies of interest to a set of ϕ_g from 0.05 steps for over the closed interval $[-0.95, 0.95]$, and the points 0.999 and 1.

The number of time periods examined are $T = \{5, 10, 20\}$, and the range of the number of cross-sections are $N = \{100, 200, 300\}$. While considered modest, these combinations are representative of dynamic panel settings that are known to exhibit poor properties, also these combinations were used in Gouriéroux, Phillips, and Yu (2010) allowing for comparison. To compute our p-value, our approach sets $H = 20$ for the simulated series and $M = 79$ simulated Monte Carlo series. The rejection frequencies are based on 1000 replications, and computed as the number times ϕ_g is rejected at a given significance level (5%) divided by 1000.

The coverage of our proposed method is comparetd with the recursive design wild bootstrap of Gonçalves and Kaffo (2014) with our implementation of their procedure outlined below. It is important to note that the bootstrap of Gonçalves and Kaffo (2014) imposes the assumption that $\sigma_\gamma = 0$, so we remain true to their settings in our implementation.

The conventional IIE asymptotic standard errors are not valid under the joint null, so coverage of these confidence intervals are not examined. This is a consequence of the breaks in the asymptotic distribution of the LSDV estimate, specifically stationary (normally distributed), local-to-unity (Ornstein-Uhlenbeck Process), and unit root (functions of Brownian motion).

5.1 Recursive-Design Wild-Bootstrap Confidence Set

Gonçalves and Kaffo (2014) introduce three distinct bootstrap approaches, but only the recursive-design residual-based wild bootstrap (RDWB) method provides good bias control, in the context of a panel-AR(1) model. Their method employs the Hahn and Kuersteiner (2002) correction term

$$\hat{\hat{\phi}} = \hat{\phi}_o + \frac{1}{T}(1 + \hat{\phi}_o) \tag{7}$$

where $\hat{\phi}_o$ is the LSDV estimate from the data. The Hahn and Kuersteiner (2002) bias-correction method has only been derived for the panel-AR(1) model, and is not appropriate for higher order autoregressive panels.

The Hahn and Kuersteiner (2002) estimate of ϕ is used to compute the estimation errors,

$$\hat{\hat{\epsilon}}_{i,t} = y_{i,t} - \hat{\hat{\mu}}_i - \hat{\hat{\phi}} y_{i,t-1},$$

for $t = 2, \ldots, T$ and for $t = 1$

$$\hat{\hat{\epsilon}}_{i,1} = y_{i,1} - \hat{\hat{\mu}}_i - \hat{\hat{\phi}} y_{i,0},$$

where

$$y_{i,0} = \frac{\hat{\hat{\mu}}_i}{1 - \hat{\hat{\phi}}}.$$

It is important to note that the initial observation, $y_{i,0}$, restricts the model to $|\phi| < 1$ under the null. As such, the RDWB method is not examined at the boundary, $\phi = 1$, in our simulation study.

The wild-bootstrap errors are generated as

$$\tilde{\epsilon}_{i,t} = \hat{\hat{\epsilon}}_{i,t} * \nu_{i,t} \text{ where } \nu_{i,t} \sim i.i.d.(0, 1).$$

In the Monte Carlo study, we use 999 independent draws of $\nu_{i,t}$ from the Rademacher distribution. The errors are used to recursively generate a set of simulated version of the dependent variable,

$$\tilde{y}_{i,t} = \hat{\mu}_i + \hat{\hat{\phi}} \tilde{y}_{i,t-1} + \tilde{\epsilon}_{i,t}.$$

Each of the bootstrapped samples are used to estimate ϕ using the bias-correction equation 7. The set of estimates are then sorted to obtain the bootstrap confidence limits of ϕ at the desired significance level.

5.2 Simulation Results

Figures 1 to 4 show the rejection frequency for the panel-AR(1) cases of the indirect confidence set inference (ICSI) method (red line) and the recursive-design wild bootstrap (RDWB) method (blue line).[3] The ICSI method has good power properties along with level control, as shown in the figures. The lowest rejection frequency coincides with the data generating process value of ϕ. As

[3]Rejection frequency tables under the null and alternatives for ICSI and RDWB are available upon request.

the magnitude of the difference between the ϕ_g and DGP value increases, the rejection frequency rises, indicating power. As T and N increase the rejection frequencies under the alternative for the ICSI method increase, while retaining level control under the null.

For $T = 5$, the RDWB method is extremely oversized for all values of ϕ. The RDWB method uses the Hahn & Keursteiner bias-correction method, which relies on Large-T. The rejection frequency improves for $T = 20$ but still well above the significance level. An increase in N, as expected, does not provide any significant improvements to size, but power would be improved if the method was correctly centered.

Table 2 show the size and power distortions that could occur when assumption 3 is not imposed and the series has been initiated in the distant past. The table shows that the model is level correct under the null ($\sigma_\gamma = 1.25$) but overrejects when a higher or lower value of σ_γ is imposed. These results corroborate the findings in table 1, where the bias distortions would lead to incorrect point estimates and poor coverage.

6 Concluding Remarks

Our discontinuous parametric definition of the initial observation leads to provable location-scale invariance of the LSDV estimate, which is suitable for joint consideration of stationary and unit root panels. Our modification results in an indirect inference objective function that is pivotal, which we invert to construct exact confidence sets. The Monte Carlo method used for our inversion works for our discontinuous joint null, where most, if not all, conventional approaches are restricted to a continuous null for confidence set construction.

In this paper, we focus on the panel-AR(1) model, following the lead of Bun and Kleibergen (2013) and Gouriéroux, Phillips, and Yu (2010), which subsumes that exogenous regressors can be partialled out.

Our simulation study assumed normally distributed errors, but our approach allows for any i.i.d. disturbances that are centered at zero and uncorrelated with the time-invariant errors. Interesting extensions include the assumption of non-Gaussian distributions like the skew-t distribution [Hansen (1994)], the generalized lambda distribution [Gorodnichenko, Mikusheva, and Ng (2012)], or the stable distribution [Beaulieu, Dufour, and Khalaf (2014)]. All of these distributions under specific parameter settings return to a Gaussian distribution as a special case.

References

Seung C Ahn and Peter Schmidt. Efficient estimation of models for dynamic panel data. *Journal of econometrics*, 68(1):5–27, 1995.

Theodore Wilbur Anderson and Cheng Hsiao. Formulation and estimation of dynamic models using panel data. *Journal of econometrics*, 18(1):47–82, 1982.

Manuel Arellano and Stephen Bond. Some tests of specification for panel data: Monte carlo evidence and an application to employment equations. *The review of economic studies*, 58(2):277–297, 1991.

Manuel Arellano and Olympia Bover. Another look at the instrumental variable estimation of error-components models. *Journal of econometrics*, 68(1):29–51, 1995.

Marie-Claude Beaulieu, Jean-Marie Dufour, and Lynda Khalaf. Multivariate tests of mean–variance efficiency with possibly non-gaussian errors. *Journal of Business & Economic Statistics*, 25(4), 2007.

Marie-Claude Beaulieu, Jean-Marie Dufour, and Lynda Khalaf. Identification-robust estimation and testing of the zero-beta capm. *The Review of Economic Studies*, 80:892–924, 2013.

Marie-Claude Beaulieu, Jean-Marie Dufour, and Lynda Khalaf. Exact confidence sets and goodness-of-fit methods for stable distributions. *Journal of Econometrics*, 181(1):3–14, 2014.

Richard Blundell and Stephen Bond. Initial conditions and moment restrictions in dynamic panel data models. *Journal of econometrics*, 87(1):115–143, 1998.

Maurice JG Bun and Martin A Carree. Bias-corrected estimation in dynamic panel data models. *Journal of Business & Economic Statistics*, 23(2):200–210, 2005.

Maurice JG Bun and Frank Kleibergen. Identification and inference in moments based analysis of linear dynamic panel data models. 2013.

Jean-Marie Dufour. Monte carlo tests with nuisance parameters: A general approach to finite-sample inference and nonstandard asymptotics. *Journal of Econometrics*, 133(2):443–477, 2006.

Jean-Marie Dufour and Pascale Valéry. Exact and asymptotic tests for possibly non-regular hypotheses on stochastic volatility models. *Journal of Econometrics*, 150(2):193–206, 2009.

A Ronald Gallant and George Tauchen. Which moments to match? *Econometric Theory*, 12(04): 657–681, 1996.

Sılvia Gonçalves and Maximilien Kaffo. Bootstrap inference for linear dynamic panel data models with individual fixed effects. *forthcoming in Journal of Econometrics*, 2014.

Yuriy Gorodnichenko, Anna Mikusheva, and Serena Ng. Estimators for persistent and possibly nonstationary data with classical properties. *Econometric Theory*, 28(05):1003–1036, 2012.

Christian Gourieroux, Alain Monfort, and Eric Renault. Indirect inference. *Journal of applied econometrics*, 8(S1):S85–S118, 1993.

Christian Gouriéroux, Peter CB Phillips, and Jun Yu. Indirect inference for dynamic panel models. *Journal of Econometrics*, 157(1):68–77, 2010.

Jinyong Hahn. How informative is the initial condition in the dynamic panel model with fixed effects? *Journal of Econometrics*, 93(2):309–326, 1999.

Jinyong Hahn and Guido Kuersteiner. Asymptotically unbiased inference for a dynamic panel model with fixed effects when both n and t are large. *Econometrica*, 70(4):1639–1657, 2002.

Jinyong Hahn and Hyungsik Roger Moon. Reducing bias of mle in a dynamic panel model. *Econometric Theory*, 22(03):499–512, 2006.

Bruce E Hansen. Autoregressive conditional density estimation. *International Economic Review*, pages 705–730, 1994.

Jan F Kiviet. On bias, inconsistency, and efficiency of various estimators in dynamic panel data models. *Journal of econometrics*, 68(1):53–78, 1995.

Jan F Kiviet and IT van den Doel. Asymptotic consequences of neglected dynamics in individual effects models*. *Statistica Neerlandica*, 48(1):71–85, 1994.

Jerzy Neyman and Elizabeth L Scott. Consistent estimates based on partially consistent observations. *Econometrica: Journal of the Econometric Society*, pages 1–32, 1948.

Stephen Nickell. Biases in dynamic models with fixed effects. *Econometrica: Journal of the Econometric Society*, pages 1417–1426, 1981.

Peter CB Phillips. On confidence intervals for autoregressive roots and predictive regression. *Econometrica*, 82(3):1177–1195, 2014.

Anthony A Smith. Estimating nonlinear time-series models using simulated vector autoregressions. *Journal of Applied Econometrics*, 8(S1):S63–S84, 1993.

A Proofs

For the proofs below we formally define the demeaned lagged dependent series as

$$g_{i,t-1} = \left(\sum_{r=0}^{t-2} \epsilon_{i,t-1-r}\phi^r - \frac{1}{T-1}\sum_{s=2}^{T}\sum_{r=0}^{s-2} \epsilon_{i,s-1-r}\phi^r\right) + \left(\sigma_\gamma \epsilon_{i,0}\phi^{t-1} - \frac{1}{T-1}\sum_{s=2}^{T} \sigma_\gamma \epsilon_{i,0}\phi^{s-1}\right),$$

which is distinct from $g_{i,t}$ since it also back-shifts the summation for the empirical average.

A.1 Proof of Theorem 1

Panel-AR(1) Scale Invariance. Let $\tilde{\epsilon}_{i,t} = \kappa\epsilon_{i,t}$ and $\tilde{\epsilon}_{i,0} = \kappa\epsilon_{i,0}$, where κ is a postive constant, leading to the transformed series $\tilde{y}_{i,t}$. The transformed demeaned series are defined as,

$$\begin{aligned}
\tilde{y}^*_{i,t} &= \tilde{y}_{i,t} - \tilde{y}_{i,\bullet} \\
&= g_i(\phi, \sigma_\gamma, \tilde{\epsilon}_{i,0}, \tilde{\epsilon}_i, t, T) = g_i(\phi, \sigma_\gamma, \kappa\epsilon_{i,0}, \kappa\epsilon_i, t, T) \\
&= \left(\sum_{r=0}^{t-1} \tilde{\epsilon}_{i,t-r}\phi^r - \frac{1}{T-1}\sum_{s=2}^{T}\sum_{r=0}^{s-1} \tilde{\epsilon}_{i,t-r}\phi^r\right) + \left(\sigma_\gamma \tilde{\epsilon}_{i,0}\phi^t - \frac{1}{T-1}\sum_{s=2}^{T} \sigma_\gamma \tilde{\epsilon}_{i,0}\phi^s\right) \\
&= \left(\sum_{r=0}^{t-1} \kappa\epsilon_{i,t-r}\phi^r - \frac{1}{T-1}\sum_{s=2}^{T}\sum_{r=0}^{s-1} \kappa\epsilon_{i,t-r}\phi^r\right) + \left(\sigma_\gamma \kappa\epsilon_{i,0}\phi^t - \frac{1}{T-1}\sum_{s=2}^{T} \sigma_\gamma \kappa\epsilon_{i,0}\phi^s\right) \\
&= \kappa\left[\left(\sum_{r=0}^{t-1} \epsilon_{i,t-r}\phi^r - \frac{1}{T-1}\sum_{s=2}^{T}\sum_{r=0}^{s-1} \epsilon_{i,t-r}\phi^r\right) + \left(\sigma_\gamma \epsilon_{i,0}\phi^t - \frac{1}{T-1}\sum_{s=2}^{T} \sigma_\gamma \epsilon_{i,0}\phi^s\right)\right] \\
&= \kappa g_i(\phi, \sigma_\gamma, \epsilon_{i,0}, \epsilon_i, t, T) \\
&= \kappa y^*_{i,t},
\end{aligned}$$

and likewise for the lagged dependent series. The series is linear in σ_ϵ, so κ is a common factor for all the terms in the series. The LSDV estimate of ϕ is

$$\begin{aligned}
\tilde{\phi} &= (\kappa y^{*\prime}_{-1}\kappa y^*_{-1})^{-1}\kappa y^{*\prime}_{-1}\kappa y^* \\
&= \frac{\kappa^2}{\kappa^2}(y^{*\prime}_{-1}y^*_{-1})^{-1}y^{*\prime}_{-1}y^* \\
&= \hat{\phi}
\end{aligned}$$

which is identical to equation 5. Therefore under assumption 1 the LSDV estimator of ϕ is scale invariant. ■

Panel-AR(1) Location Invariance. Let $\tilde{\mu}_i$ represent a transformed set of random coefficients, where the transformation can be either a level shift, $\tilde{\mu}_i = \delta_i + \mu_i$, or relative shift, $\tilde{\mu}_i = \kappa_i \mu_i$, or both, $\tilde{\mu}_i = \delta_i + \kappa_i \mu_i$. The new demeaned series, $\tilde{y}_{i,t}$, follows as above leading to

$$\begin{aligned}
\tilde{y}^*_{i,t} &= \tilde{y}_{i,t} - \tilde{y}_{i,\bullet} \\
&= \left[\tilde{\mu}_i + \left(\sum_{r=0}^{t-1} \epsilon_{i,t-r}\phi^r\right) + \sigma_\gamma \epsilon_{i,0}\phi^t\right] - \left[\tilde{\mu}_i + \left(\frac{1}{T-1}\sum_{s=2}^{T}\sum_{r=0}^{s-1} \epsilon_{i,t-r}\phi^r\right) + \frac{1}{T-1}\sum_{s=2}^{T} \sigma_\gamma \epsilon_{i,0}\phi^s\right] \\
&= \left(\sum_{r=0}^{t-1} \epsilon_{i,t-r}\phi^r - \frac{1}{T-1}\sum_{s=2}^{T}\sum_{r=0}^{s-1} \epsilon_{i,t-r}\phi^r\right) + \left(\sigma_\gamma \epsilon_{i,0}\phi^t - \frac{1}{T-1}\sum_{s=2}^{T} \sigma_\gamma \epsilon_{i,0}\phi^s\right) \\
&= g_i(\phi, \sigma_\gamma, \epsilon_{i,0}, \epsilon_i, t, T) \\
&= y^*_{i,t},
\end{aligned}$$

and likewise for the lagged dependent series. The demeaned series are independent of $\tilde{\mu}$ and equal to the demeaned series if $\delta_i \neq 0$ or $\kappa_i = 1$. Hence, the LSDV estimate of ϕ from the transformed data is identical to the untransformed data. Therefore under assumptions 1 the LSDV estimator of ϕ is location invariant. ■

A.2 Corollary Proofs

Proof of Corollary 1: Affine Transformation Invariance. The LSDV estimate is invariant to affine transformations of the data, consider the transformation $\tilde{y}_{i,t} = (\kappa y_{i,t} + \delta_i)$. The transformed series can be defined as,

$$\begin{aligned}
\tilde{y}_{i,t} &= \kappa\mu_i + \kappa \sum_{r=0}^{t-1} \epsilon_{i,t-r}\phi^r + \kappa\sigma_\gamma\epsilon_{i,0}\phi^t + \delta_i \\
&= (\kappa\mu_i + \delta_i) + \kappa\left(\sum_{r=0}^{t-1} \epsilon_{i,t-r}\phi^r + \sigma_\gamma\epsilon_{i,0}\phi^t\right).
\end{aligned}$$

The empirical average is given by

$$\tilde{y}_{i,\bullet} = (\kappa\mu_i + \delta_i) + \kappa\frac{1}{T-1}\sum_{s=2}^{T}\left(\sum_{r=0}^{s-1}\epsilon_{i,t-r}\phi^r + \sigma_\gamma\epsilon_{i,0}\phi^s\right),$$

which leads to the demeaned series,

$$\begin{aligned}
\tilde{y}^*_{i,t} &= \tilde{y}_{i,t} - \tilde{y}_{i,\bullet} \\
&= \kappa\left(\sum_{r=0}^{t-1}\epsilon_{i,t-r}\phi^r + \sigma_\gamma\epsilon_{i,0}\phi^t\right) - \kappa\frac{1}{T-1}\sum_{s=2}^{T}\left(\sum_{r=0}^{s-1}\epsilon_{i,t-r}\phi^r + \sigma_\gamma\epsilon_{i,0}\phi^s\right) \\
&= \kappa g_i(\phi, \sigma_\gamma, \epsilon_{i,0}, \epsilon_i, t, T) \\
&= \kappa y^*_{i,t},
\end{aligned}$$

and likewise for the lagged dependent series. By scale invariance of the estimator, we find that $\tilde{\phi} = \hat{\phi}$, so the LSDV estimate is invariant to affine transformations. ■

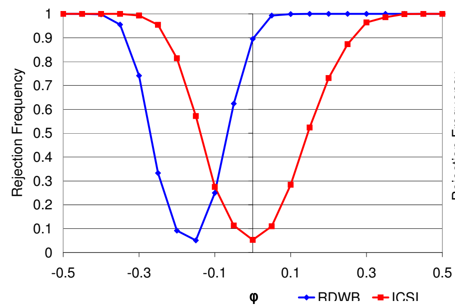
Proof of Corollary 2: Invariance to σ_γ at Unity. The LSDV estimate is invariant to σ_γ for $\phi = 1$. Consider the demeaned series under unity,

$$\begin{aligned}
y^*_{i,t}|_{\phi=1} &= g_i(\phi, \sigma_\gamma, \epsilon_{i,0}, \epsilon_i, t, T|\phi = 1) \\
&= \left(\sum_{r=0}^{t-1}\epsilon_{i,t-r} - \frac{1}{T-1}\sum_{s=2}^{T}\sum_{r=0}^{s-1}\epsilon_{i,t-r}\right) + \left(\sigma_\gamma\epsilon_{i,0} - \frac{1}{T-1}\sum_{s=2}^{T}\sigma_\gamma\epsilon_{i,0}\right) \\
&= \left(\sum_{r=0}^{t-1}\epsilon_{i,t-r} - \frac{1}{T-1}\sum_{s=2}^{T}\sum_{r=0}^{s-1}\epsilon_{i,t-r}\right) + (\sigma_\gamma\epsilon_{i,0} - \sigma_\gamma\epsilon_{i,0}) \\
&= \left(\sum_{r=0}^{t-1}\epsilon_{i,t-r} - \frac{1}{T-1}\sum_{s=2}^{T}\sum_{r=0}^{s-1}\epsilon_{i,t-r}\right),
\end{aligned}$$

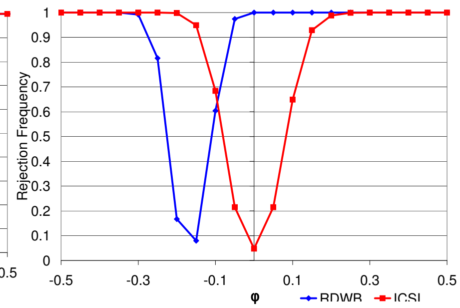
and likewise for the lagged dependent series. Therefore the LSDV estimate of ϕ is not a function of σ_γ, as such simulated series under the unit root null allow for any choice of σ_γ. By setting $\sigma_\gamma = 0$ when $\phi = 1$, we recover the conditional setting outlined in assumption 3. ■

B Figures

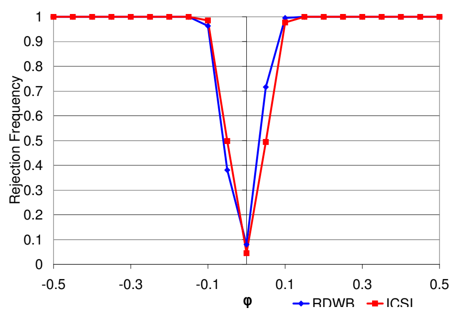
Figure 1: Rejection Frequency for $\alpha = 5\%$ and $\phi = 0$: ICSI and RDWB



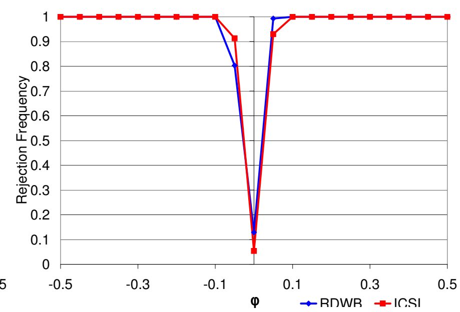
(a) N=100,T=5



(b) N=300,T=5



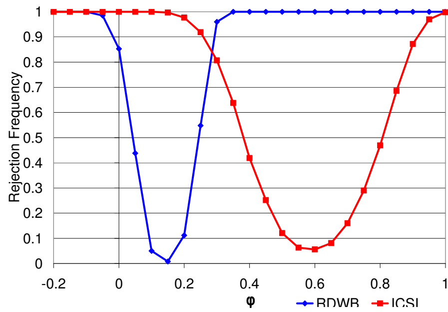
(c) N=100,T=20



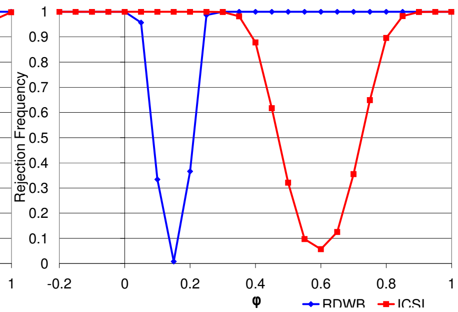
(d) N=300,T=20

Rejection frequencies based on 1000 replications. The wild bootstrap is implemented using 999 draws from the Rademacher distribution. ICSI simulations use 79 Monte Carlo replications to compute the p-value, and 20 simulations for the indirect inference computations.

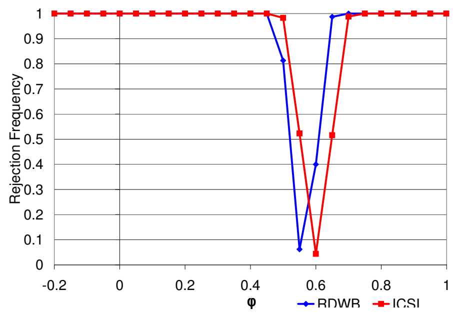
Figure 2: Rejection Frequency for $\alpha = 5\%$ and $\phi = 0.6$: ICSI and RDWB



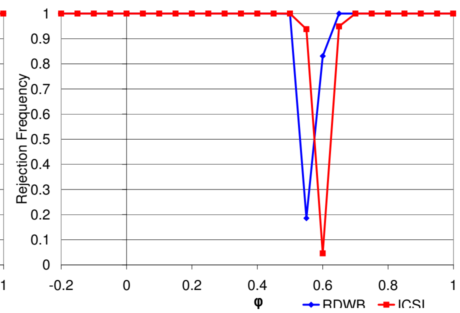
(a) N=100,T=5



(b) N=300,T=5



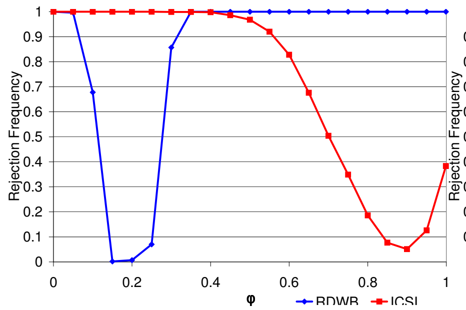
(c) N=100,T=20



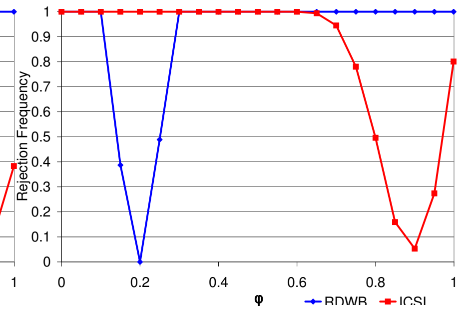
(d) N=300,T=20

Rejection frequencies based on 1000 replications. The wild bootstrap is implemented using 999 draws from the Rademacher distribution. ICSI simulations use 79 Monte Carlo replications to compute the p-value, and 20 simulations for the indirect inference computations.

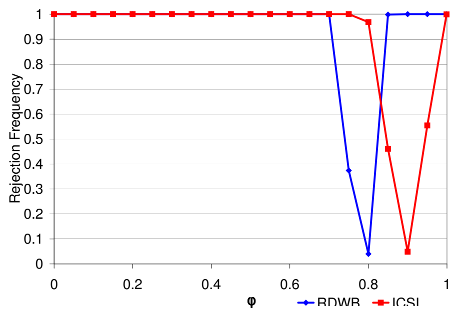
Figure 3: Rejection Frequency for $\alpha = 5\%$ and $\phi = 0.9$: ICSI and RDWB



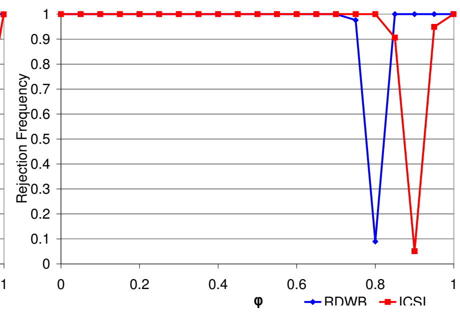
(a) N=100,T=5



(b) N=300,T=5



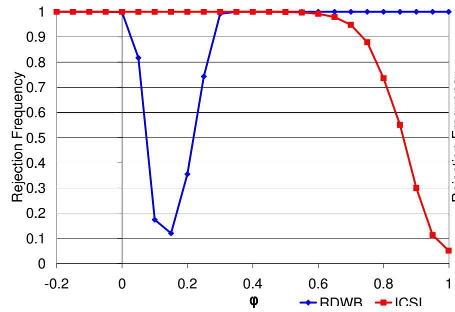
(c) N=100,T=20



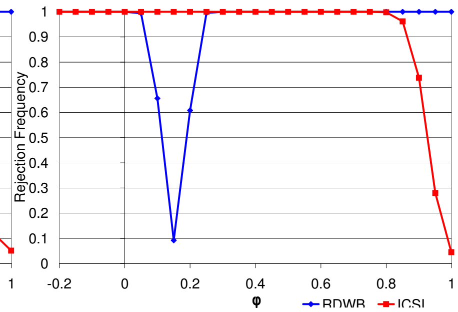
(d) N=300,T=20

Rejection frequencies based on 1000 replications. The wild bootstrap is implemented using 999 draws from the Rademacher distribution. ICSI simulations use 79 Monte Carlo replications to compute the p-value, and 20 simulations for the indirect inference computations.

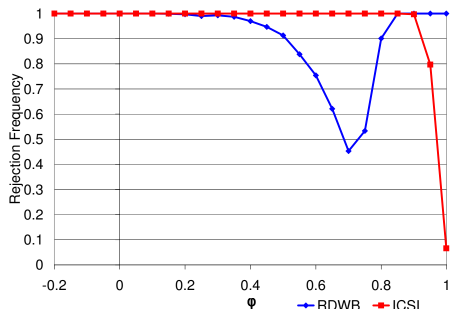
Figure 4: Rejection Frequency for $\alpha = 5\%$ and $\phi = 0.999$: ICSI and RDWB



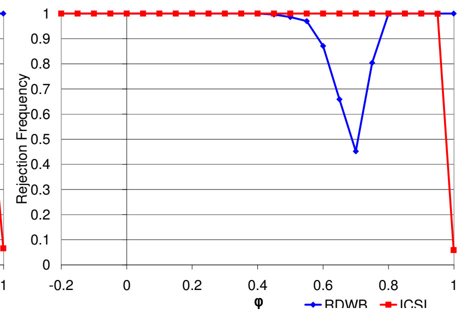
(a) N=100,T=5



(b) N=300,T=5



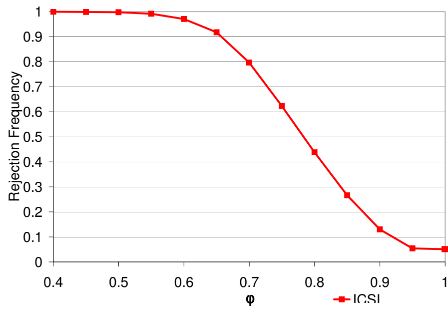
(c) N=100,T=20



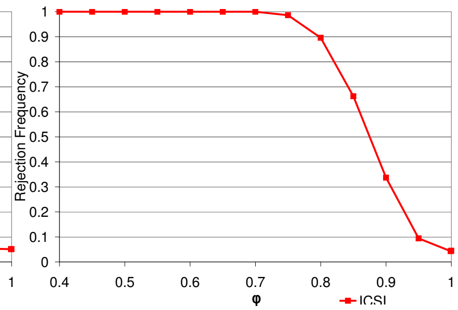
(d) N=300,T=20

Rejection frequencies based on 1000 replications. The wild bootstrap is implemented using 999 draws from the Rademacher distribution. ICSI simulations use 79 Monte Carlo replications to compute the p-value, and 20 simulations for the indirect inference computations.

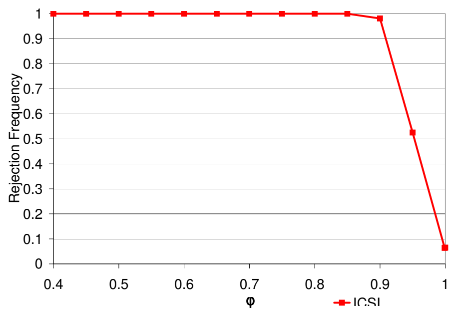
Figure 5: Rejection Frequency for $\alpha = 5\%$ and $\phi = 1.0$: ICSI



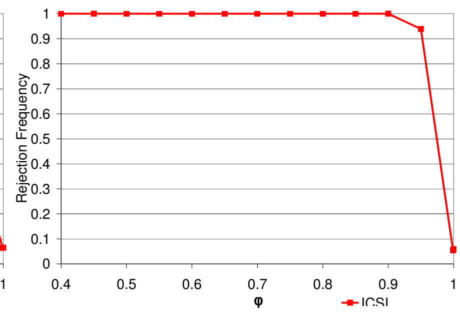
(a) N=100,T=5



(b) N=300,T=5



(c) N=100,T=20



(d) N=300,T=20

Rejection frequencies based on 1000 replications. ICSI simulations use 79 Monte Carlo replications to compute the p-value, and 20 simulations for the indirect inference computations.

Table 2: Rejection Frequency Sensitivity to Choice of σ_γ

	$\sigma_{\gamma,g}$					
ϕ_g	0	1	1.25	1.5	2	3
0.3	0.999	0.999	0.999	0.999	0.994	0.965
0.35	0.988	0.982	0.975	0.966	0.925	0.714
0.4	0.939	0.896	0.870	0.815	0.660	0.301
0.45	0.794	0.655	0.591	0.507	0.314	0.064
0.5	0.557	0.374	0.297	0.206	0.080	0.165
0.55	0.309	0.155	0.092	0.055	0.060	0.551
0.6	0.129	0.044	0.048	0.059	0.202	0.830
0.65	0.039	0.051	0.091	0.167	0.489	0.958
0.7	0.046	0.146	0.238	0.393	0.709	0.990
0.75	0.153	0.353	0.470	0.604	0.865	0.999
0.8	0.410	0.592	0.680	0.784	0.935	1.000

Indirect Confidence Set Inference method based on 1000 Monte Carlo replications, $N = 200$, $T = 5$, $\phi = 0.6$ and 5% significance level.