Econometric Theory, **25**, 2009, 1348–1391. Printed in the United States of America.
doi:10.1017/S0266466608090531

GMM ESTIMATION AND INFERENCE IN DYNAMIC PANEL DATA MODELS WITH PERSISTENT DATA

HUGO KRUINIGER
Queen Mary, University of London

In this paper we consider generalized method of moments–based (GMM-based) estimation and inference for the panel AR(1) model when the data are persistent and the time dimension of the panel is fixed. We find that the nature of the weak instruments problem of the Arellano–Bond (Arellano and Bond, 1991, *Review of Economic Studies* 58, 277–297) estimator depends on the distributional properties of the initial observations. Subsequently, we derive local asymptotic approximations to the finite-sample distributions of the Arellano–Bond estimator and the System estimator, respectively, under a variety of distributional assumptions about the initial observations and discuss the implications of the results we obtain for doing inference. We also propose two Lagrange multiplier–type (LM-type) panel unit root tests.

1. INTRODUCTION

In this paper we consider generalized method of moments–based (GMM-based) estimation and inference for the panel AR(1) model $y_{i,t} = \rho y_{i,t-1} + (1-\rho)\mu_i + \varepsilon_{i,t}$, $i = 1, \ldots, N$ and $t = -S+1, \ldots, T$, when the autoregression (AR) parameter ρ is close to or equal to one. Throughout the paper we assume that $S \geq -1$, the first observations occur at $t = 1$, the time dimension of the panel, T, is fixed, and the cross-section dimension of the panel, N, is large. Among other things we derive local asymptotic approximations to the finite-sample distributions of some well-known linear GMM estimators for this model under a variety of assumptions about the initial observations and discuss the implications of the results we obtain for doing inference. We also propose two Lagrange multiplier-type (LM–type) panel unit root tests.

Persistent data raise at least three issues in connection with GMM estimation of the panel AR(1) model. First, some of the available moment conditions do not identify the unit root. For instance, it is well known that the GMM estimator of Arellano and Bond (1991) (the estimator is henceforth referred to as the AB

I thank Steve Bond and Frank Windmeijer for kindly making one of their computer programs available to me. I also thank Liudas Giraitis, three anonymous referees, a co-editor, and the editor for very helpful comments. This research was funded by the ESRC under grant R000239139. Address correspondence to Hugo Kruiniger, Department of Economics, Queen Mary, University of London, Mile End Road, London E1 4NS, United Kingdom; e-mail: h.kruiniger@qmul.ac.uk.

 © 2009 Cambridge University Press 0266-4666/09 $15.00

https://doi.org/10.1017/S0266466608090531 Published online by Cambridge University Press

estimator) breaks down when $\rho = 1$, because the instruments,which are lagged levels of the data ($y_{i,s}$, $s \leq t-2$), are no longer correlated with the first-differences of the regressors ($\Delta y_{i,t-1}$): when $\rho = 1$, $\mathrm{E}(y_{i,s}\Delta y_{i,t-1}) = 0$. However, moment conditions such as $\mathrm{E}[\Delta y_{i,s}(y_{i,t} - \rho y_{i,t-1})] = 0$ with $s \leq t-1$ still identify the unit root; see Arellano and Bover (1995) (henceforth Arbov). In what follows we will refer to an estimator that only exploits such moment conditions as an Arbov estimator. Second, most moment conditions are weak in some sense when ρ is close to unity. The finite-sample distributions of the corresponding estimators may therefore differ substantially from the first-order fixed-parameter asymptotic distributions. The second-order bias approximation for the AB estimator derived by Hahn, Hausman, and Kuersteiner (2001) also becomes inaccurate when ρ is close to unity. Third, the first-order derivatives with respect to ρ of some of the moment conditions are discontinuous at $\rho = 1$: it is easily verified for $t \geq 2$ that $\lim_{\rho\uparrow 1}\mathrm{E}(y_{i,1}\Delta y_{i,t}) = -\frac{1}{2}\mathrm{Var}(\varepsilon_{i,t}) \equiv -\frac{1}{2}\sigma^2$ when the data are covariance stationary, whereas $\mathrm{E}(y_{i,1}\Delta y_{i,t}) = 0$ when $\rho = 1$.

Recently, Han and Phillips (2007) proposed estimators for the panel AR(1) model that avoid the weak moment conditions problem when ρ is near unity. However the validity of their strong moment conditions relies on covariance stationarity of the initial conditions. Phillips and Han (2008) discuss a related least squares estimator for the time series AR(1) model.

In the paper it is shown that the weakness of the moment conditions implied by the panel AR(1) model can manifest itself in different ways depending on the distributional properties of the initial observations. In some cases the first-order derivatives of the moment conditions with respect to ρ, i.e., the Jacobians, tend to zero when ρ approaches one. In other cases the moment conditions are weak because the standard deviations of the first-order derivatives of the underlying moment functions with respect to ρ explode when ρ approaches one. In the first type of situation the "signal" of the moment functions becomes weak, whereas in the second type of situation the overwhelming "noise" of the moment functions drowns their signal when ρ approaches one (cf. Han and Phillips, 2006). In both situations doing a form of local asymptotics affords a better approximation to the finite-sample distribution of the corresponding GMM estimator than doing conventional first-order fixed-parameter asymptotics.

When S is fixed and ρ approaches one according to $\rho = 1 - \lambda N^{-1/2}$ as N grows large, where $\lambda > 0$, both the means and the standard deviations of the cross-sectional averages of the first-order derivatives of the AB moment functions with respect to ρ vanish at a $\sqrt{N}$ rate. Thus when S is small one can obtain local-to-nonidentification approximations to the finite-sample distributions of AB estimators by applying the local-to-zero asymptotics of Staiger and Stock (1997). On the other hand, when the data are covariance stationary, the standard deviations of the first-order derivatives of the AB moment functions with respect to ρ explode when ρ approaches one, whereas their means tend to a nonzero constant because of the fact that $\mathrm{Var}(y_{i,t} - \mu_i) = \sigma^2/(1-\rho^2)$. In this case one can obtain local-to-nonidentification asymptotic approximations

https://doi.org/10.1017/S0266466608090531 Published online by Cambridge University Press

to the finite-sample distributions of AB estimators by assuming that ρ approaches one according to $\rho = 1 - \lambda N^{-1}$ as N grows large, where again $\lambda > 0$. Furthermore, we show that in this case the Arbov estimators and the two-step "optimal" System estimator (cf. the paper by Arellano and Bover, 1995) have nonnormal distributions also under such local asymptotics although they are still consistent.

The cases of fixed S and covariance stationarity, which corresponds to $S = \infty$, represent two extreme cases for the local-to-nonidentification asymptotic approximation to the finite-sample distributions of the AB estimator. Furthermore, under covariance stationarity and $\rho = 1 - \lambda N^{-1}$, we have $\mathrm{Var}(y_{i,t}) = O(N)$, which implies that the variance of the data is large when N is large. This implication may not be compatible with the data in hand. Therefore we also derive in this paper for all $d \geq 0$ the local-to-nonidentification asymptotic distributions of the AB estimator under the assumption that both S and N grow large with $S/N^d \to c^* > 0$. We show that if ρ approaches one according to $\rho = 1 - \lambda N^{-g}$ as N grows large, where $\lambda > 0$ and $0 < g \leq 1$, and if $0 \leq d \leq g$, then $\mathrm{Var}(y_{i,t}) = O(N^d)$. Note that the variance of the data in hand can be used to choose a suitable value for d. Moreover, in this setup $\mathrm{E}(y_{i,s}\Delta y_{i,t-1}) = O(N^{d-g})$ and $[\mathrm{Var}(y_{i,s}\Delta y_{i,t-1})]^{1/2} = O(N^{(1/2)d})$, where $s \leq t-2$. This means that both the strength of the signal and the noise of the AB moment functions increase with d, i.e., with S. The value of g for which the local-to-nonidentification asymptotic distributions of AB estimators are obtained also increases with d from a value of $1/2$ for $d = 0$ to a maximum value of 1 for $d \geq 1$. Thus a larger value of S strengthens the identification of ρ by the AB moment conditions. The value of g is chosen in such a way that the means and the standard deviations of the cross-sectional averages of the first-order derivatives of the AB moment functions with respect to ρ either do not vanish or vanish at the same rate when ρ approaches one according to $\rho = 1 - \lambda N^{-g}$ as N grows large, where $\lambda > 0$. When $d > 1$ and $g = 1$, both the Arbov and the two-step "optimal" System estimator have the same nonnormal local asymptotic distributions as under covariance stationarity when $g = 1$.

In the paper we also show that under local (-to-nonidentification) asymptotics (1) the "optimal" AB estimators are biased downward, (2) the estimators of the optimal weight matrices for the Arbov estimator and the System estimator and their asymptotic standard errors are no longer consistent when $g = 1$ and the data are covariance stationary or $d \geq 1$, (3) the vector of cross-sectional averages of the Arbov moment functions and the vector of their first-order derivatives with respect to ρ are uncorrelated, (4) the two-step "optimal" System estimator can have a skewed distribution, and (5) the truly optimal Arbov and System estimators have the same distribution, which is symmetric. The first result explains the downward biases of the AB estimator reported in a Monte Carlo study contained in Blundell and Bond (1998). The other results help to explain the fact that when the data are covariance stationary and persistent, Wald tests based on the two-step "optimal" System estimator have incorrect size, whereas LM tests that

https://doi.org/10.1017/S0266466608090531 Published online by Cambridge University Press

are based on System estimators that use a restricted estimator of the weight matrix that is optimal under the null have correct size. This paper therefore offers a theoretical justification for using LM tests in the context of panel AR models. Finally the paper proposes two LM-type panel unit root tests and studies their properties.

The paper is organized as follows. In Section 2 we review GMM estimation of the panel AR(1) model. In Section 3 we derive local asymptotic approximations to the finite-sample distributions of the AB, Arbov, and System estimators under various assumptions about the initial conditions, i.e., under various asymptotic plans for S and N. Section 4 proposes two LM-type panel unit root tests and investigates their properties both analytically and through Monte Carlo experiments. Section 5 concludes. Proofs are given in the Appendix.

A few words on notation. We use the symbol $\xrightarrow{d}$ to signify convergence in distribution, the symbol plim to signify convergence in probability, and the symbol $\xrightarrow{q.m.}$ to signify convergence in quadratic mean. To state multi-index asymptotic results we make use of the following notation (for definitions of the underlying concepts, see also Phillips and Moon, 1999): $\lim_{S,N\to\infty,seq} N^{-1}\sum_{i=1}^{N} X_{i,t} = X_t$ is equivalent to $\lim_{N\to\infty}\lim_{S\to\infty} N^{-1}\sum_{i=1}^{N} X_{i,t} = X_t$ whereas $N^{-1/2}\sum_{i=1}^{N} X_{i,t} \xrightarrow{d} X_t$ as $(S, N\to\infty)_{seq}$ signifies that $N^{-1/2}\sum_{i=1}^{N} X_{i,t}$ converges in distribution to X_t sequentially by letting S pass to infinity first and letting N pass to infinity subsequently. Sometimes we only write $\lim_{N\to\infty} N^{-1}\sum_{i=1}^{N} X_{i,t} = X_t$ instead of $\lim_{S,N\to\infty,seq} N^{-1}\sum_{i=1}^{N} X_{i,t} = X_t$ when it is clear that S passes to infinity first, e.g., when we have assumed covariance stationarity. Finally, $\operatorname{plim}_{N\to\infty,S/N\to c} N^{-1}\sum_{i=1}^{N} X_{i,t} = X_t$ is an example of a so-called diagonal path probability limit. In this example both S and N pass to infinity with S/N converging to some constant c. We also make use of indicator functions. For instance, $1\{d = 0\} = 1$ if $d = 0$ and $1\{d = 0\} = 0$ if $d \neq 0$. Finally, the abbreviation PD denotes positive definite.

2. GMM ESTIMATION OF THE PANEL AR(1) MODEL

2.1. The Panel AR(1) Model

Consider the panel AR(1) model with random effects:

$$y_{i,t} = \rho y_{i,t-1} + v_{i,t}, \tag{1}$$

$$v_{i,t} = \eta_i + \varepsilon_{i,t}, \quad \text{where } \eta_i = (1-\rho)\mu_i, \tag{2}$$

for $i = 1,\ldots,N$ and $t = -S+1,\ldots,T$. The starting date of the $\{y_{i,t}\}$, $t = -S$, need not coincide with the date of the initial observations on y, $t = 1$, i.e., $-S \leq 1$; the number of "individuals," N, is large whereas the number of observations on y per individual, T, is fixed. Moreover $-1 < \rho \leq 1$.

https://doi.org/10.1017/S0266466608090531 Published online by Cambridge University Press

We assume random sampling in the cross section.[1] The composite error terms, the $v_{i,t}$, satisfy the following standard assumptions (cf. Ahn and Schmidt, 1995):[2]

$$\mathrm{E}(\mu_i) = 0 \quad \text{and} \quad \mathrm{E}(\mu_i^2) = \sigma_\mu^2 < \infty, \tag{3}$$

$$\mathrm{E}(\varepsilon_{i,t}) = 0, \qquad \mathrm{E}(\varepsilon_{i,t}\eta_i) = 0, \quad \text{and} \tag{4}$$

$$\mathrm{E}(\varepsilon_{i,t}\, y_{i,-S}) = 0, \quad \text{for } t = -S+1, \ldots, T, \tag{5}$$

and

$$\mathrm{E}(\varepsilon_{i,s}\varepsilon_{i,t}) = 0, \quad \text{for } s \neq t, \qquad s, t = -S+1, \ldots, T. \tag{6}$$

For convenience we also assume that the idiosyncratic errors, the $\varepsilon_{i,t}$, are homoskedastic over time:

$$\mathrm{E}(\varepsilon_{i,t}^2) = \sigma^2 > 0, \quad \text{for } t = -S+1, \ldots, T. \tag{7}$$

However, in the paper we do not consider any moment conditions based on this assumption. The initial conditions are given by $y_{i,-S} = \mu_i + (1-\rho)\xi_{i,-S}$, for $i = 1, \ldots, N$. Note that when $\rho = 1$, the initial conditions are equal to the individual effects, the μ_i. Finally, we assume that $\mathrm{E}(\xi_{i,-S}^8) < \infty$, $\mathrm{E}(\mu_i^4) < \infty$, $\mathrm{E}(\varepsilon_{i,t}^4) = \kappa\sigma^4$, and $\mathrm{E}(\varepsilon_{i,t}^8) < \infty$, for $t = -S+1, \ldots, T$, where κ is the kurtosis parameter. We also define $\sigma_{-S}^2 = \mathrm{E}(y_{i,-S}^2)$.

Note that the model can be rewritten as $y_{i,t} - \mu_i = \rho(y_{i,t-1} - \mu_i) + \varepsilon_{i,t}$ and that the η_i disappear (the μ_i drop out) when $\rho = 1$. The parametrization $\eta_i = (1-\rho)\mu_i$ prevents the individual effects from turning into individual trends at $\rho = 1$ and thereby avoids a discontinuity in the data generating process at $\rho = 1$ (cf. Han and Phillips, 2007). Furthermore, when combined with the assumption $\sigma_\mu^2 < \infty$, these restrictions on the η_i ensure that under covariance stationarity the means of the data, i.e., $\eta_i/(1-\rho)$, $i = 1, \ldots, N$, are drawn from a distribution with a finite variance rather than a variance that tends to infinity when ρ approaches one. Without these restrictions, some of our asymptotic results would be different and/or become discontinuous at $\rho = 1$.

In the paper we consider various versions of the panel AR(1) model that differ with respect to the assumptions made about the initial observations. Among them are the following two versions:

(CS) The $\{y_{i,t}\}$ have reached covariance stationarity at $t = 1$ when $|\rho| < 1$.

(FS) Fixed S: thus the $\{y_{i,t}\}$ have not reached stationarity at $t = 1$ when $|\rho| < 1$.

Ahn and Schmidt (1997) have shown that given assumptions (3)–(7), the $\{y_{i,t}\}$ have reached covariance stationarity at $t = 1$ if and only if the initial observations satisfy the following assumptions:

$$\mathrm{E}(y_{i,1} - \mu_i) = 0, \qquad \mathrm{E}[(y_{i,1} - \mu_i)\mu_i] = 0, \quad \text{and} \tag{8}$$

$$\mathrm{Var}(y_{i,1} - \mu_i) = \frac{\sigma^2}{1-\rho^2}, \quad \text{when } |\rho| < 1. \tag{9}$$

https://doi.org/10.1017/S0266466608090531 Published online by Cambridge University Press

Note that the FS model does not rule out that mean-stationarity, i.e., assumption (8), holds.

2.2. GMM Estimators for the Panel AR(1) Model

Arellano and Bond (1991) proposed a GMM estimator that exploits the following $m = (T-1)(T-2)/2$ linear moment conditions:

$$\mathrm{E}[m_{AB,s,t}(\rho)] = \mathrm{E}[y_{i,s}(\Delta y_{i,t} - \rho \Delta y_{i,t-1})] = 0,$$
$$\text{for } s = 1, \ldots, t-2 \quad \text{and} \quad t = 3, \ldots, T, \tag{10}$$

where $m_{AB,s,t}(\rho) = y_{i,s}(\Delta y_{i,t} - \rho \Delta y_{i,t-1})$ and $\Delta y_{i,t} = y_{i,t} - y_{i,t-1}$. These moment conditions are implied by assumptions (4)–(6). Note that they do not identify the unit root because $\mathrm{E}(y_{i,s} \Delta y_{i,t-1}) = \mathrm{E}(y_{i,s} \varepsilon_{i,t-1}) = 0$ when $\rho = 1$.

Arellano and Bover (1995) noted that if mean-stationarity, i.e., assumption (8), holds also, one can add $T-2$ linear moment conditions to those in (10):

$$\mathrm{E}[(y_{i,t} - \rho y_{i,t-1}) \Delta y_{i,t-1}] = 0, \quad \text{for } t = 3, \ldots, T. \tag{11}$$

The latter moment conditions do identify the unit root because $\mathrm{E}(y_{i,t-1} \Delta y_{i,t-1}) = \mathrm{E}(y_{i,t-1} \varepsilon_{i,t-1}) = \sigma^2$ when $\rho = 1$.

A GMM estimator that exploits the moment conditions in both (10) and (11) is known as a System (SYS) estimator.

The set of moment conditions in (10) and (11) is equivalent to a set that contains $T-2$ AB-type and m Arbov-type moment conditions:

$$\mathrm{E}[y_{i,1}(\Delta y_{i,t} - \rho \Delta y_{i,t-1})] = 0, \quad \text{for } t = 3, \ldots, T, \tag{12}$$

and

$$\mathrm{E}[m_{Arbov,s,t}(\rho)] = \mathrm{E}[(y_{i,t} - \rho y_{i,t-1}) \Delta y_{i,s}] = 0,$$
$$\text{for } s = 2, \ldots, t-1 \quad \text{and} \quad t = 3, \ldots, T, \tag{13}$$

where $m_{Arbov,s,t}(\rho) = (y_{i,t} - \rho y_{i,t-1}) \Delta y_{i,s}$. A GMM estimator that only exploits the latter m moment conditions will be referred to as an Arbov estimator.

Let $y_i^t = [y_{i,1} \ldots y_{i,t}]$ and let $Z_i = \mathrm{diag}(y_i^1, \ldots, y_i^{T-2})$ be a $(T-2) \times m$ block-diagonal matrix. Then we can write the set of AB moment conditions in (10) as $\mathrm{E}(Z_i' \Delta v_i) = 0$ where $\Delta v_i = v_i - v_{i,-1} = [\Delta v_{i,3} \ \ldots \ \Delta v_{i,T}]'$. Under our assumptions, $\mathrm{E}(\Delta v_i \Delta v_i')/\sigma^2 = H$, where $H = H_{T-2}$ is a $(T-2)$ band-diagonal matrix with 2's on the main diagonal, -1's on the first sub- and superdiagonal, and zeros elsewhere. It follows that the AB GMM estimator that uses $W_{N,AB1} = (N^{-1} \sum_{i=1}^N Z_i' H Z_i)^{-1}$ as weight matrix is an optimal one-step GMM estimator. This estimator is denoted as $\widehat{\rho}_{AB1}$. In this paper $\{W_N\}$ denotes an arbitrary sequence of PD weight matrices with $\mathrm{plim}_{N \to \infty} W_N = W$, where W is PD. An AB estimator based on W_N is simply denoted as $\widehat{\rho}_{AB}$.

https://doi.org/10.1017/S0266466608090531 Published online by Cambridge University Press

There exist no feasible optimal one-step weight matrices for the Arbov and System estimators, except when $\sigma_\eta^2 = 0$. Let $\Delta y_i^t = [\Delta y_{i,2} \ \dots \ \Delta y_{i,t}]$, let $Z_i^I = y_{i,1} I_{T-2}$, where I_{T-2} is an $(T-2)$ identity matrix, let $Z_i^{II} = \text{diag}(\Delta y_i^2, \dots, \Delta y_i^{T-1})$ be a $(T-2) \times m$ block-diagonal matrix, and let $Z_i^S = \text{diag}(Z_i^I, Z_i^{II})$ be a $2(T-2) \times (T-2+m)$ block-diagonal matrix. When $\sigma_\eta^2 = 0$, optimal one-step weight matrices for the Arbov and System estimators are given by $W_{N,Arbov1} = (N^{-1} \sum_{i=1}^N Z_i^{II\prime} Z_i^{II})^{-1}$ and $W_{N,SYS1} = (N^{-1} \sum_{i=1}^N Z_i^{S\prime} A Z_i^S)^{-1}$, respectively, where A is given by

$$A = \begin{bmatrix} H & C \\ C' & I_{T-2} \end{bmatrix},$$

where $C = C_{T-2}$ is a $(T-2) \times (T-2)$ matrix with ones on the main diagonal, -1's on the first subdiagonal, and zeros elsewhere. Note that $C = \text{E}(v_i \Delta v_i')'/\sigma^2$. The one-step GMM estimators based on $W_{N,Arbov1}$ and $W_{N,SYS1}$ will be referred to as the Arbov1 estimator and the SYS1 estimator, respectively. Finally, let $W_{N,SYS1b} = (N^{-1} \sum_{i=1}^N Z_i^{S\prime} Z_i^S)^{-1}$.

The "optimal" two-step System (SYS2) estimator is based on the weight matrix $W_{N,SYS2}(\widehat{\rho}_I) = (N^{-1} \sum_{i=1}^N Z_i^{S\prime} [\widehat{\Delta v_i'} \ \widehat{v_i'}]' [\widehat{\Delta v_i'} \ \widehat{v_i'}] Z_i^S)^{-1}$, with $\widehat{\Delta v}_i = \widehat{\Delta v}_i(\widehat{\rho}_I) = \Delta y_i - \widehat{\rho}_I \Delta y_{i,-1}$, $\widehat{v}_i = y_i - \widehat{\rho}_I y_{i,-1}$, and, unless stated otherwise, $\widehat{\rho}_I = \widehat{\rho}_1$, where $\widehat{\rho}_1$ is an initial $\sqrt{N}$-consistent estimator for ρ, i.e., $\sqrt{N}(\widehat{\rho}_1 - \rho) \xrightarrow{d} \mathcal{K} \neq 0$. We define the "optimal" two-step AB (AB2) and Arbov (Arbov2) estimators and their weight matrices analogously. The $W_{N,SYSk}$ matrices $(k = 1, 2)$ can be partitioned as

$$W_{N,SYSk} = \begin{bmatrix} W_{N,SYSk,11} & W_{N,SYSk,12} \\ W_{N,SYSk,21} & W_{N,SYSk,22} \end{bmatrix},$$

where the $W_{N,SYSk,11}$ block corresponds to $\text{E}(Z_i^{I\prime} \Delta v_i) = 0$.

Let the $(T-2) \times (T-t-2)$ matrix d_t be given by $d_t = [0 \ I_{T-t-2}]'$. In addition, let $Z_i^{AB} = [\, y_{i,1} I_{T-2} \ Z_i^D] = [Z_i^I \ Z_i^D]$, where $Z_i^D = [d_1 \Delta y_{i,2} \ d_2 \Delta y_{i,3} \ \dots \ d_{T-3} \Delta y_{i,T-2}]$ is a $(T-2) \times (m - (T-2))$ matrix. Then we can rewrite the set of m AB moment conditions in $\text{E}(Z_i' \Delta v_i) = 0$ as $\text{E}(Z_i^{AB\prime} \Delta v_i) = 0$. Thus there exists a nonsingular constant matrix K^{AB} such that $Z_i' = K^{AB} Z_i^{AB\prime}$.

An alternative transformation that can be used for removing individual effects is the Helmert transformation, which amounts to taking forward orthogonal deviations. The Helmert transformation of $(v_{i,2} \ v_i')'$ is $\widetilde{v}_i = (\widetilde{v}_{i,2} \dots \widetilde{v}_{i,T-1})'$, where $\widetilde{v}_{i,t} = ((T-t)/(T-t+1))^{1/2} \left[v_{i,t} - 1/(T-t) \sum_{s=t+1}^T v_{i,s} \right]$, $t = 2, \dots, T-1$. Note that $v_{i,t} - 1/(T-t) \sum_{s=t+1}^T v_{i,s}$ is equal to a linear combination of first differences of the errors. An advantage of using the Helmert transformation rather than taking first differences is that it preserves the orthogonality among the errors; i.e., if $\text{E}(\varepsilon_i \varepsilon_i') = \sigma^2 I_{T-2}$, then $\text{E}(\widetilde{\varepsilon}_i \widetilde{\varepsilon}_i') = \sigma^2 I_{T-3}$.

https://doi.org/10.1017/S0266466608090531 Published online by Cambridge University Press

Applying the Helmert transformation to $y_{i,t} = \rho y_{i,t-1} + (1-\rho)\mu_i + \varepsilon_{i,t}$, $t = 2, \ldots, T$, yields $\widetilde{y}_{i,t} = \rho \widetilde{y}_{i,-1,t-1} + \widetilde{\varepsilon}_{i,t}$, $t = 2, \ldots, T-1$, where $\widetilde{y}_{i,-1,t-1} = ((T-t)/(T-t+1))^{1/2}\left[y_{i,t-1} - 1/(T-t)\sum_{s=t}^{T-1} y_{i,s}\right]$. Note that all lagged values of $y_{i,t}$ are valid instruments for the tth equation of the transformed system. If we let $z_{i,t}' = y_i^{t-1}$, then the set of AB moment conditions in (10) is equivalent to the following set of moment conditions: $\mathrm{E}(z_{i,t}\widetilde{v}_{i,t}) = 0$, $t = 2, \ldots, T-1$. That is, there is a nonsingular matrix K^H such that $Z_i'\Delta v_i = K^H Z_i'\widetilde{v}_i$; see Arellano and Bover (1995). An optimal weight matrix for $\mathrm{E}(Z_i'\widetilde{v}_i) = 0$ is $[N^{-1}\sum_{i=1}^{N}(Z_i'Z_i)]^{-1}$, which is a block-diagonal matrix with typical diagonal block equal to $[N^{-1}\sum_{i=1}^{N}(z_{i,t}z_{i,t}')]^{-1}$. Note that $Z_i'HZ_i = K^H Z_i'Z_i K^{H\prime}$. It follows that the one-step AB GMM estimator can be rewritten as a weighted average of $T-2$ two-stage least squares estimators: $\widehat{\rho}_{AB1} = \{\sum_{t=2}^{T-1}[\widetilde{y}_{-1,t-1}'Z_t(Z_t'Z_t)^{-1}Z_t'\widetilde{y}_{-1,t-1}]\}^{-1}$ $\sum_{t=2}^{T-1}[\widetilde{y}_{-1,t-1}'Z_t(Z_t'Z_t)^{-1}Z_t'\widetilde{y}_t]$ where $\widetilde{y}_t = [\widetilde{y}_{1,t} \ldots \widetilde{y}_{N,t}]'$, $\widetilde{y}_{-1,t-1} = [\widetilde{y}_{1,-1,t-1} \ldots \widetilde{y}_{N,-1,t-1}]'$, and $Z_t = [z_{1,t} \ldots z_{N,t}]'$.

3. ASYMPTOTIC PROPERTIES OF THE ESTIMATORS WHEN THE DATA ARE PERSISTENT

Here we investigate the asymptotic properties of various AB, Arbov, and System estimators when ρ is close to unity under two different asymptotic schemes: $S, N \to \infty$ sequentially with $S \to \infty$ first (the CS model) and $S, N \to \infty$ simultaneously with $S/N^d \to c^* > 0$, where $d \geq 0$. The latter scheme includes $N \to \infty$, and S is fixed (the FS model) as special case for $d = 0$. Note that the model assumes that T is fixed.

We will see that the estimators suffer from some kind of weak instruments problem when ρ is close to unity, with the exact nature of the problem depending on the distributional properties of the initial conditions. Thus ρ is not only the parameter of interest but also determines the "strength" of the instruments. Conventional fixed parameter asymptotics yields a poor approximation to the finite-sample distribution of the estimators when ρ is close to unity because of the fact that the weak instruments problem vanishes when N increases. To ensure that the weak instruments problem is captured by the asymptotics, we let ρ tend to unity as a function of N when $N \to \infty$ in such a way that asymptotically ρ is local-to-nonidentified. That is, our asymptotics will make use of parameter sequences such as $\rho = 1 - \lambda N^{-g}$ where $\lambda > 0$ and $\frac{1}{2} \leq g \leq 1$. The rate N^g at which ρ tends to unity depends on the distributional properties of the initial conditions.

Note that the parameter sequences $\rho = 1 - \lambda N^{-g}$ are introduced to obtain better asymptotic approximations to the finite-sample distribution of the estimators for ρ when the instruments are weak and not because of interest in local-to-unit-root models. Thus the motivation for considering these parameter sequences is different from that in the time series literature. Moon and Phillips (2000) have considered consistent panel data estimators for the localizing parameter $\gamma < 0$ in $\rho = \exp(\gamma/T)$ assuming that T grows large.

https://doi.org/10.1017/S0266466608090531 Published online by Cambridge University Press

In the analysis the ratios $\mathrm{Var}(y_{i,t}-\mu_i)/\sigma^2$ play an important role. The panel AR(1) model implies the following expression for $\mathrm{Var}(y_{i,1}-\mu_i)$ when $|\rho|<1$:

$$\mathrm{Var}(y_{i,1}-\mu_i)=\rho^{2(1+S)}\mathrm{Var}(y_{i,-S}-\mu_i)+(1-\rho^{2(1+S)})\sigma^2/(1-\rho^2). \tag{14}$$

It is easily verified for any S and $\rho>0$ that $d[\mathrm{Var}(y_{i,1}-\mu_i)]/d\rho>0$ and $\lim_{\rho\uparrow 1}\mathrm{Var}(y_{i,1}-\mu_i)=\mathrm{Var}(y_{i,-S}-\mu_i)+(1+S)\sigma^2$. Moreover, $\lim_{S\to\infty}\mathrm{Var}(y_{i,1}-\mu_i)=\sigma^2/(1-\rho^2)$. It follows that when S is large and ρ tends to one, the ratio $\mathrm{Var}(y_{i,1}-\mu_i)/\sigma^2$ becomes large. Indeed, $\lim_{\rho\uparrow 1}\lim_{S\to\infty}\mathrm{Var}(y_{i,1}-\mu_i)=\infty$. However, when S is fixed $\lim_{\rho\uparrow 1}\mathrm{Var}(y_{i,1}-\mu_i)<\infty$.

3.1. Local Asymptotics for the Covariance Stationary Case

We assume that the CS version of the model applies, i.e., $S\to\infty$ and the model has reached stationarity at $t=1$ when $|\rho|<1$. In this case the finite-sample distributions of the AB and System GMM estimators differ considerably from their first-order fixed-parameter asymptotic distributions when ρ is close to one; see, e.g., the Monte Carlo evidence reported in Blundell and Bond (1998). To understand why this happens, we first examine the underlying moment conditions.

When the CS model holds, $\lim_{\rho\uparrow 1}\mathrm{Var}(y_{i,t})/\sigma^2=\lim_{\rho\uparrow 1}[\mathrm{Var}(y_{i,t}-\mu_i)/\sigma^2+\sigma_\mu^2/\sigma^2]=\lim_{\rho\uparrow 1}[1/(1-\rho^2)]+\sigma_\mu^2/\sigma^2=\infty$ for $t=1,\ldots,T$. Furthermore, $\mathrm{E}[(\Delta y_{i,t-1})^2]=2\sigma^2/(1+\rho)$ and $\mathrm{E}(-y_{i,s}\Delta y_{i,t-1})=(1-\rho)\rho^{t-2-s}\mathrm{E}[(y_{i,s}-\mu_i+\mu_i)(y_{i,s}-\mu_i)]=\sigma^2(1-\rho)\rho^{t-2-s}/(1-\rho^2)$ when $s\le t-2$. The latter equalities imply that $\lim_{\rho\uparrow 1}\mathrm{E}(dm_{AB,s,t}/d\rho)=\lim_{\rho\uparrow 1}\mathrm{E}(-y_{i,s}\Delta y_{i,t-1})=\sigma^2/2$ for $s\le t-2$ and $\lim_{\rho\uparrow 1}\mathrm{E}(dm_{Arbov,s,t}/d\rho)=\lim_{\rho\uparrow 1}\mathrm{E}(-y_{i,t-1}\Delta y_{i,s})=-\sigma^2/2$ when $s\le t-1$.

Consider now the simple first-stage regression $\Delta y_{i,t-1}=\pi y_{i,s}+\omega_i$ $(s\le t-2)$, which corresponds to a typical AB moment condition. Then we find that $\lim_{\rho\uparrow 1}\mathrm{plim}_{N\to\infty}\widehat{\pi}=\lim_{\rho\uparrow 1}\mathrm{plim}_{N\to\infty}(\sum_i y_{i,s}^2)^{-1}\sum_i y_{i,s}\Delta y_{i,t-1}=0$. Furthermore, if $\rho=1-\lambda N^{-1}$, the first-stage F-statistic $F_{AB}=\widehat{\pi}^2N(\sum_i y_{i,s}^2)/\sum_i(\Delta y_{i,t-1}-\widehat{\pi}y_{i,s})^2=O_p(1)$ and $\pi=[\mathrm{E}(y_{i,s}^2)]^{-1}\mathrm{E}(y_{i,s}\Delta y_{i,t-1})=O(N^{-1})$. The expression $F_{AB}-1$ provides an estimate of the so-called concentration parameter, which is a measure of the strength of the instrument(s); see, e.g., Stock, Wright, and Yogo (2002) for details. The fact that $F_{AB}=O_p(1)$ suggests that there is a potential weak instruments problem. Note that if one would choose the parameter sequence $\rho=1-\lambda N^{-1/2}$, then $\pi=O(N^{-1/2})$ but $F_{AB}=O_p(N^{1/2})$. Thus when the CS version of the model applies and ρ is sufficiently close to unity, the AB GMM estimator suffers from some sort of weak instruments problem, albeit not from one of the Staiger and Stock (1997) type where $F_{AB}=O_p(1)$ when $\pi=O(N^{-1/2})$. In this case the weak instruments problem arises because $\lim_{\rho\uparrow 1}\mathrm{E}(y_{i,s}^2)\to\infty$, whereas $\lim_{\rho\uparrow 1}\mathrm{E}(y_{i,s}\Delta y_{i,t-1})=-\sigma^2/2\neq 0$. Considering the multiple first-stage regressions $\Delta y_{i,t-1}=\sum_{k=1}^{t-2}\pi_k y_{i,k}+\omega_i$ for $t=3,\ldots,T$ leads to the same conclusion because T is fixed.

https://doi.org/10.1017/S0266466608090531 Published online by Cambridge University Press

Next consider the simple first-stage regression $y_{i,t-1} = \pi \Delta y_{i,s} + \omega_i$ $(s \leq t-1)$, which corresponds to a typical Arbov moment condition. Clearly, $\lim_{\rho \uparrow 1} \operatorname{plim}_{N\to\infty} \widehat{\pi} = 1/2 \neq 0$. Nonetheless, if $\rho = 1 - \lambda N^{-1}$, the first-stage F-statistic $F_{Arbov} = \widehat{\pi}^2 N \sum_i (\Delta y_{i,s})^2 / \sum_i (y_{i,t-1} - \widehat{\pi} \Delta y_{i,s})^2 = O_p(1)$, even though $\pi = O(1)$. The reason for this finding is that when $S \to \infty$, $\mathrm{E}(y_{i,t-1}^2) \propto 1/(1-\rho)$ and hence $\operatorname{plim}_{N\to\infty} N^{-2} \sum_i (y_{i,t-1} - \widehat{\pi} \Delta y_{i,s})^2 = \sigma^2/(2\lambda)$ if $\rho = 1 - \lambda N^{-1}$. However, $\lim_{\rho \uparrow 1} \mathrm{E}(y_{i,t-1} \Delta y_{i,s}) = \sigma^2/2 \neq 0$.

The preceding analysis suggests that when the CS model holds and ρ approaches one the finite-sample distributions of the AB, Arbov, and System estimators are affected by exploding higher order moments and that one should not use the local-to-zero asymptotics of Staiger and Stock (1997) to obtain approximations to these distributions. Under covariance stationarity not only $\mathrm{Var}(y_{i,1}) \propto 1/(1-\rho)$ but also $\mathrm{Var}(y_{i,1} \Delta y_{i,2}) \propto 1/(1-\rho)$, $\mathrm{Var}(y_{i,2} \Delta y_{i,2}) \propto 1/(1-\rho)$, $\mathrm{Var}(y_{i,1} \Delta \varepsilon_{i,3}) \propto 1/(1-\rho)$, and $\mathrm{Cov}(y_{i,1} \Delta \varepsilon_{i,3},\ y_{i,1} \Delta y_{i,2}) \propto 1/(1-\rho)$; see Lemma A.2 in the Appendix. In this case one can obtain local (-to-nonidentification) approximations to the distributions of the estimators by choosing a parameter sequence such that the variances of the sample covariances of the instruments and the differenced regressors become $O(1)$, i.e., by choosing $\rho = 1 - \lambda N^{-1}$. Under this sequence $\mathrm{Var}(y_{i,1}) = O(N)$. A large value of $\mathrm{Var}(y_{i,1})$ may or may not be an accurate description of the data.[3]

We have the following local (-to-nonidentification) asymptotic results for the AB and Arbov estimators.

THEOREM 1. *Let the CS model hold, $T = 3$, and $\rho = 1 - \lambda N^{-1}$ with $\lambda > 0$. Then, as $N \to \infty$,*

$$\textit{(i)}\ \widehat{\rho}_{AB} - \rho \xrightarrow{d} \frac{\tilde{X}_1}{\tilde{X}_2}, \quad \textit{where} \begin{bmatrix} \tilde{X}_1 \\ \tilde{X}_2 \end{bmatrix} \sim N\left[-\tfrac{1}{2}\sqrt{\lambda} \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \tfrac{1}{2} \begin{pmatrix} 2 & -1 \\ -1 & 1 \end{pmatrix} \right].$$

$$\textit{(ii)}\ \sqrt{N}(\widehat{\rho}_{Arbov} - \rho) \xrightarrow{d} \frac{\tilde{X}_3}{\tilde{X}_4}, \quad \textit{where} \begin{bmatrix} \tilde{X}_3 \\ \tilde{X}_4 \end{bmatrix} \sim N\left[\tfrac{1}{2} \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & \frac{1}{2\lambda} \end{pmatrix} \right].$$

The distribution of the ratio of two, possibly correlated, normal variates does not have finite moments; see, e.g., Fieller (1932). Furthermore, the density of such a ratio may be unimodal or bimodal. Noting that such a ratio can be rewritten as $c_1 + c_2(a+w)/(b+x)$ where $a \geq 0$, $b \geq 0$, c_1 and c_2 are constants, and w and x are independent standard normal variates, Marsaglia (1965) presents a necessary and sufficient condition for bimodality in terms of the values of a and b; see Figure 2 in his paper. For the Arbov estimator in Theorem 1(ii) we have $a = 0$, and this confirms that its limit density is unimodal. For the AB estimator in Theorem 1(i) we obtain $a = b = \sqrt{\lambda/2}$, and its limit density is bimodal if and only if $\lambda > \lambda_0$ with $\lambda_0 \approx 4$. Even if the second mode exists, it is not of practical importance: as pointed out by Woglom (2001), a second mode of practical importance requires that the value of b is substantially less than the value of a. For instance, when

https://doi.org/10.1017/S0266466608090531 Published online by Cambridge University Press

$a = b = 2.25$, the density at the second mode is only about 0.2% of the density at the primary mode. A recent discussion of the possible bimodality of the (limit) density of instrumental variables (IV) estimators can be found in the papers by Forchini (2006), Hillier (2006), and Phillips (2006).

The local asymptotic results of Theorem 1 have been obtained under the assumption of covariance stationarity. However, if we take $\lambda = 0$ in the local asymptotic distribution of the AB estimator, we obtain its distribution for $\rho = 1$, despite the fact that we need to condition on the initial conditions in this case.[4] Let $\rho = 1$ and $T = 3$. Then it can easily be shown that as $(S, N \to \infty)_{seq}$,[5]

$$\widehat{\rho}_{AB} - \rho \xrightarrow{d} \frac{\tilde{X}_1}{\tilde{X}_2}, \quad \text{where } \begin{bmatrix} \tilde{X}_1 \\ \tilde{X}_2 \end{bmatrix} \sim N\left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 & -1 \\ -1 & 1 \end{pmatrix} \right] \tag{15}$$

and

$$\sqrt{S}(\widehat{\rho}_{Arbov} - \rho) \xrightarrow{d} \frac{\tilde{X}_3}{\tilde{X}_4}, \quad \text{where } \begin{bmatrix} \tilde{X}_3 \\ \tilde{X}_4 \end{bmatrix} \sim N\left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \right]. \tag{16}$$

The Arbov estimator remains consistent under local asymptotics, but its local asymptotic distribution is nonnormal when $\rho = 1 - \lambda N^{-g}$ with $g \geq 1$. When $g < 1$, its local distribution is equal to its first-order fixed-parameter asymptotic distribution, whereas if $g \geq 1$, $N^{(1/2)g}(\widehat{\rho}_{Arbov} - \rho) \xrightarrow{d} \tilde{X}_3/\tilde{X}_{4g}$ with $\tilde{X}_{4g} \sim N(1\{g = 1\}(1/2), 1/(2\lambda))$ and $\mathrm{Cov}(\tilde{X}_3, \tilde{X}_{4g}) = 0$.

The set of m Arbov moment conditions in $\mathrm{E}(Z_i^{II\prime} v_i) = 0$ can be restated as $\mathrm{E}(\tilde{Z}_i^{II\prime}[\Delta v_i' \; v_i']') = 0$, where $\tilde{Z}_i^{II} = \mathrm{diag}(Z_i^D, Z_i^L)$ is a $2(T-2) \times m$ matrix with $Z_i^L = \mathrm{diag}(\Delta y_{i,2}, \ldots, \Delta y_{i,T-1})$. Let K^{II} be the nonsingular constant matrix such that $\tilde{Z}_i^{II\prime}[\Delta v_i' \; v_i']' = K^{II} Z_i^{II\prime} v_i$. Then we can generalize Theorem 1 to $T \geq 3$ using the following lemma.

LEMMA 1. *Let the CS model hold, $T \geq 3$, and $\rho = 1 - \lambda N^{-1}$ with $\lambda > 0$. Let $N \to \infty$. Then*

(i) (1) $N^{-1} \sum_{i=1}^N Z_i^{I\prime} \Delta v_i \xrightarrow{d} \tilde{X}_{51}$,
$N^{-1} \sum_{i=1}^N Z_i^{I\prime} \Delta y_{i,-1} \xrightarrow{d} \tilde{X}_{61}$,
$N^{-1/2} \sum_{i=1}^N Z_i^{D\prime} \Delta v_i \xrightarrow{d} \tilde{X}_{52}$, *and*
$N^{-1/2} \sum_{i=1}^N Z_i^{D\prime} \Delta y_{i,-1} \xrightarrow{d} \tilde{X}_{62}$ *with*
$\tilde{X}_5 = (\tilde{X}_{51}' \; \tilde{X}_{52}')' \sim N(0, \tilde{\Sigma}_{55})$, $\tilde{X}_6 = (\tilde{X}_{61}' \; \tilde{X}_{62}')' \sim N(\tilde{\mu}_6, \tilde{\Sigma}_{66})$, *and* $\mathrm{E}(\tilde{X}_5 \tilde{X}_6') = \tilde{\Sigma}_{56} \neq 0$, *where $\tilde{\mu}_6$, $\tilde{\Sigma}_{55}$, $\tilde{\Sigma}_{56}$, and $\tilde{\Sigma}_{66}$ are given in the proof;*

(2) when $T = 3$, $\tilde{X}_5 = \tilde{X}_{51} = \sigma^2 \tilde{X}_1/\sqrt{\lambda}$ and $\tilde{X}_6 = \tilde{X}_{61} = \sigma^2 \tilde{X}_2/\sqrt{\lambda}$, where $\tilde{X}_1$ and $\tilde{X}_2$ are defined in Theorem 1.

(ii) (1) $N^{-1/2} \sum_{i=1}^N Z_i^{L\prime} v_i \xrightarrow{d} \tilde{X}_{71}$ *and*
$N^{-1} \sum_{i=1}^N Z_i^{L\prime} y_{i,-1} \xrightarrow{d} \tilde{X}_{81} \sim N((\sigma^2 \iota/2), (\sigma^4/(2\lambda))I)$ *with*

https://doi.org/10.1017/S0266466608090531 Published online by Cambridge University Press

$\tilde{X}_7 = (\tilde{X}_{52}' \ \tilde{X}_{71}')' \sim N(0, \tilde{\Sigma}_{77})$, $\tilde{\Sigma}_{77} = \sigma^4 K^{II} K^{II\prime}$, *and* $\mathrm{E}(\tilde{X}_7 \tilde{X}_8') = 0$, *where* $\tilde{X}_8 = (0' \ \tilde{X}_{81}')'$;

(2) *when* $T = 3$, $\tilde{X}_7 = \tilde{X}_{71} = \sigma^2 \tilde{X}_3$ *and* $\tilde{X}_8 = \tilde{X}_{81} = \sigma^2 \tilde{X}_4$, *where* $\tilde{X}_3$ *and* $\tilde{X}_4$ *are defined in Theorem 1.*

In addition, we have the following results concerning the weight matrices:

(iii) $\mathrm{plim}_{N\to\infty} N^{-2} \sum_{i=1}^N (Z_i^{I\prime} Z_i^I) = (\sigma^2/(2\lambda)) I$,
$\mathrm{plim}_{N\to\infty} N^{-3/2} \sum_{i=1}^N (Z_i^{I\prime} Z_i^D) = 0$, *and*
$\mathrm{plim}_{N\to\infty} N^{-1} \sum_{i=1}^N (Z_i^{D\prime} Z_i^D) = \sigma^2 I$.

(iv) $\mathrm{plim}_{N\to\infty} W_{N,Arbov1} = \sigma^{-2} I$ *and*

$$N^{-1} \sum_{i=1}^N \left(\tilde{Z}_i^{II\prime} [\widehat{\Delta v_i'} \ \widehat{v_i'}]' [\widehat{\Delta v_i'} \ \widehat{v_i'}] \tilde{Z}_i^{II} \right) \xrightarrow{d} \sigma^4 \tilde{W}_{Arbov2}^{-1}$$

$$\equiv \sigma^4 K^{II} K^{II\prime} + (\sigma^4/(2\lambda)) \mathcal{K}^2 \mathrm{diag}(O_{m-(T-2)}, I_{T-2}).$$

THEOREM 2. *Let the CS model hold,* $T \geq 3$, *and* $\rho = 1 - \lambda N^{-1}$ *with* $\lambda > 0$. *Let* $N \to \infty$. *Then*

(i) (1) $\widehat{\rho}_{AB1} - \rho \xrightarrow{d} \tilde{X}_6' \tilde{\Sigma}_{55}^{-1} \tilde{X}_5 / \tilde{X}_6' \tilde{\Sigma}_{55}^{-1} \tilde{X}_6$. *If* $T > 3$, $\widehat{\rho}_{AB1}$ *is asymptotically biased downward.*

(2) *assuming that* $\widehat{\rho}_{AB}$ *exploits* $\mathrm{E}(Z_i^{AB\prime} \Delta v_i) = 0$ *in lieu of* $\mathrm{E}(Z_i' \Delta v_i) = 0$,

$$\widehat{\rho}_{AB} - \rho \xrightarrow{d} \tilde{X}_{61}' W_{11} \tilde{X}_{51} / \tilde{X}_{61}' W_{11} \tilde{X}_{61}.$$

(ii) $\mathrm{plim}_{N\to\infty} (\widehat{\rho}_{Arbov} - \rho) = 0$, *and assuming that* $\widehat{\rho}_{Arbov}$ *exploits* $\mathrm{E}(\tilde{Z}_i^{II\prime} \times [\Delta v_i' \ v_i']') = 0$ *in lieu of* $\mathrm{E}(Z_i^{II\prime} v_i) = 0$,

$$\sqrt{N}(\widehat{\rho}_{Arbov} - \rho) \xrightarrow{d} \tilde{X}_8' W \tilde{X}_7 / \tilde{X}_8' W \tilde{X}_8.$$

Parts (ii) of Lemma 1 and Theorem 2 imply that because $\tilde{X}_8$ is Gaussian and hence $\tilde{X}_7 \perp \tilde{X}_8$, i.e., $\tilde{X}_7$ and $\tilde{X}_8$ are independent, the correct local asymptotic standard errors of Arbov estimators are given by $(\mathrm{E}(\tilde{X}_8' W \tilde{\Sigma}_{77} W \tilde{X}_8 / (\tilde{X}_8' W \tilde{X}_8)^2))^{0.5}$. Thus when the data are CS the conventional asymptotic standard errors of Arbov estimators are inconsistent under $\rho = 1 - \lambda N^{-1}$. Moreover, because the truly local asymptotically optimal weight matrix is proportional to $\tilde{\Sigma}_{77}^{-1}$, the second result in Lemma 1(iv) implies that the conventional estimator of the optimal weight matrix for the Arbov estimator is also inconsistent under these local asymptotics.[6] However, the Arbov estimator that uses $W_{N,Arbov2}(\widehat{\rho}_1)$ as weight matrix is still consistent, although when $T > 3$ its local asymptotic distribution is different from that of the truly optimal Arbov estimator. Finally, if $W \perp (\tilde{X}_7' \ \tilde{X}_8')'$ and $\rho = 1 - \lambda N^{-1}$, the local asymptotic distribution of the Arbov estimator is symmetric and, provided $T > 3$, the Arbov estimator is unbiased under these local asymptotics. In

https://doi.org/10.1017/S0266466608090531 Published online by Cambridge University Press

contrast, when the data are CS and $\rho = 1 - \lambda N^{-1}$, $\mathrm{E}(\tilde{X}_5 \tilde{X}_6') \neq 0$; i.e., the cross-products of the instruments and the regressors in the AB estimator remain correlated with the cross-products of the instruments and the dependent variables when $N \to \infty$, thereby causing the bias of this estimator.

Note that the set of $m+T-2$ System moment conditions in $\mathrm{E}(Z_i^{S\prime}[\Delta v_i' \; v_i']') = 0$ can be rewritten as $\mathrm{E}(\tilde{Z}_i^{S\prime}[\Delta v_i' \; v_i']') = 0$, where $\tilde{Z}_i^S = \mathrm{diag}(Z_i^{AB}, \; Z_i^L)$ is a $2(T-2) \times (m+T-2)$ matrix. We have the following results for the System estimator.

THEOREM 3. *Let the CS model hold, let $T \geq 3$, and let $\rho = 1 - \lambda N^{-1}$ with $\lambda > 0$. Let $N \to \infty$. Let $\tilde{W}_{SYS*} = \mathrm{diag}(O_{T-2}, \; I_m)$, where O_{T-2} is a $(T-2) \times (T-2)$ null matrix. Finally, let $D = \mathrm{E}[\mathrm{diag}(\varepsilon_i^2, \ldots, \varepsilon_i^{T-1})' \iota \Delta \varepsilon_i'] / \sigma^2$. Then*

(i) $\mathrm{plim}_{N\to\infty} W_{N,SYS1} = \sigma^{-2} \tilde{W}_{SYS*}$ *and* $\mathrm{plim}_{N\to\infty} W_{N,SYS1b} = \sigma^{-2} \tilde{W}_{SYS*}$.

(ii) $\mathrm{diag}(N^{-1} I_{T-2}, \; N^{-1/2} I_m) \sum_{i=1}^N \big(\tilde{Z}_i^{S\prime}[\widehat{\Delta v_i'} \; \widehat{v_i'}]' [\widehat{\Delta v_i'} \; \widehat{v_i'}] \tilde{Z}_i^S\big) \mathrm{diag}(N^{-1} I_{T-2}, \; N^{-1/2} I_m) \xrightarrow{d} \sigma^4 \tilde{W}_{SYS2}^{-1}$ *with* $\tilde{W}_{SYS2}^{11} \equiv (\frac{1}{2\lambda}) H$, $\tilde{W}_{SYS2}^{21} = (\tilde{W}_{SYS2}^{12})' \equiv -(\frac{1}{2\lambda}) \mathcal{K} K^{II} D$, *and* $\tilde{W}_{SYS2}^{22} \equiv \tilde{W}_{Arbov2}^{-1}$, *where* $\tilde{W}_{Arbov2}^{-1}$ *is defined in Lemma 1.*

(iii) $\mathrm{plim}_{N\to\infty}(\widehat{\rho}_{SYSk} - \rho) = 0$, *for* $k = 1, 1b, 2$, *and*

$$\sqrt{N}(\widehat{\rho}_{SYSk} - \rho) \xrightarrow{d} \tilde{X}_8' \tilde{\Sigma}_{77}^{-1} \tilde{X}_7 / \tilde{X}_8' \tilde{\Sigma}_{77}^{-1} \tilde{X}_8, \quad \text{for } k = 1, 1b,$$

$$\sqrt{N}(\widehat{\rho}_{SYS2} - \rho) \xrightarrow{d} \tilde{X}_{11}' \tilde{W}_{SYS2} \tilde{X}_{10} / \tilde{X}_{11}' \tilde{W}_{SYS2} \tilde{X}_{11}$$

with $\tilde{X}_{10} = (\tilde{X}_{51}' \; \tilde{X}_7')'$ *and* $\tilde{X}_{11} = (0 \; \tilde{X}_8')'$, *where* $\tilde{\Sigma}_{77}$, $\tilde{X}_{51}$, $\tilde{X}_7$, *and* $\tilde{X}_8$ *are defined in Lemma 1.*

(iv) assuming that $\widehat{\rho}_{SYS}$ *exploits* $\mathrm{E}(\tilde{Z}_i^{S\prime}[\Delta v_i' \; v_i']') = 0$ *in lieu of* $\mathrm{E}(Z_i^{S\prime}[\Delta v_i' \; v_i']') = 0$,

$$\widehat{\rho}_{SYS} - \rho \xrightarrow{d} (\tilde{X}_{61}' W_{11} \tilde{X}_{51} + \tilde{X}_{81}' W_{31} \tilde{X}_{51}) / \tilde{X}_{12}' W \tilde{X}_{12}$$

with $\tilde{X}_{12} = (\tilde{X}_{61}' \; 0 \; \tilde{X}_{81}')'$, *where* $\tilde{X}_{51}$, $\tilde{X}_{61}$, *and* $\tilde{X}_{81}$ *are defined in Lemma 1.*

Part (iv) of Theorem 3 implies that when the data are CS, a System estimator that uses a weight matrix estimator that has a PD probability limit is inconsistent under $\rho = 1 - \lambda N^{-1}$. Moreover, the conventional estimator for the optimal weight matrix for the System estimator, i.e., $W_{N,SYS2}(\widehat{\rho}_1)$, and the conventional asymptotic standard errors of System estimators are also inconsistent under such asymptotics. Nevertheless, the System estimators that use $W_{N,SYS1}$ and $W_{N,SYS2}(\widehat{\rho}_1)$, respectively, as weight matrix are still consistent under such asymptotics although the local asymptotic distribution of the two-step System estimator is asymmetric because $\mathrm{E}(\tilde{X}_{51} \tilde{X}_{81}') \neq 0$. The truly optimal System estimator has the same local asymptotic distribution as the truly optimal Arbov

https://doi.org/10.1017/S0266466608090531 Published online by Cambridge University Press

estimator, which is symmetric. Thus the local asymptotic distribution of the truly optimal System estimator is not affected by the AB moment conditions.

The preceding results suggest that Wald tests based on the usual two-step System estimator will have incorrect size when the data are CS and ρ is close to one as a consequence of its asymmetric distribution. On the other hand, LM tests that are based on Arbov or System estimators that use a restricted estimator of the weight matrix that is optimal under the null, e.g., $W_{N,Arbov2}(\rho)$ and $W_{N,SYS2}(\rho)$, respectively, have a standard normal asymptotic distribution under the null both when first-order fixed-parameter asymptotics and when local asymptotics are employed because of the fact that $\tilde{X}_7 \perp \tilde{X}_8$.[7] Therefore, such LM tests will have rejection probabilities very close to their nominal size.

3.2. Diagonal Path Local Asymptotics

The results in Section 3.1 are largely based on the fact that if $\rho = 1 - \lambda N^{-1}$, then the sequential limits $\lim_{N\to\infty}\lim_{S\to\infty}[\mathrm{Var}(y_{i,t} - \mu_i)]/N = \sigma^2/2\lambda > 0$, for $t = 1, 2, \ldots, T$. Now suppose that $\rho = 1 - \lambda N^{-1}$ and that $N \to \infty$, $S \to \infty$ simultaneously with $S/N \to c > 0$, where c is a constant; then $\lim_{N\to\infty, S/N\to c}\rho^{2(1+S)} = \exp(-2\lambda c) < 1$ because $\lim_{p\to\infty}(1 + x/p)^p = \exp(x)$. Using (14) it follows that $\lim_{N\to\infty, S/N\to c}[\mathrm{Var}(y_{i,t} - \mu_i)]/N = \sigma^2[1 - \exp(-2\lambda c)]/2\lambda > 0$. This result suggests that when S is of the same order of magnitude as N and the data are persistent, then the AB and Arbov moment conditions are still weak. Moreover, in this case one can still obtain approximations to the distributions of the AB, Arbov, and SYS estimators by using local asymptotics.

It turns out that we can derive diagonal path local asymptotic results for N, $S \to \infty$ with $(S/N^d) \to c^* > 0$ for any $d \geq 0$. Let $q(\lambda, c) = 1 - \exp(-2\lambda c)$. Then we have the following diagonal path local asymptotic results for the AB and Arbov estimators when $T = 3$.

THEOREM 4. *Let $T = 3$ and let $\rho = 1 - \lambda N^{-g}$ with $\lambda > 0$ and $0 < g \leq 1$. Let $0 \leq d \leq g$ and $\Delta_{gd} = 1/2(1-d) - (g-d)$. Furthermore, let $N \to \infty$, $S \to \infty$ simultaneously with $S/N \to c \geq 0$, $S/N^g \to c^+ \geq 0$, and $S/N^d \to c^* > 0$. Let $\hat{q}(\lambda, c^*) = 2\lambda(S+1)$ if $d = 0$, $\hat{q}(\lambda, c^*) = 2\lambda c^*$ if $0 < d < g$, and $\hat{q}(\lambda, c^*) = q(\lambda, c^*)$ if $d = g$. Finally, let $\sigma_{\mu y} = \mathrm{E}(y_{i,1}\mu_i)$ and $\sigma_y^2 = \mathrm{E}(y_{i,1}^2)$. Then*

(i)

$$N^{\Delta_{gd}}(\widehat{\rho}_{AB} - \rho) \xrightarrow{d} N(0, 2\sigma^2\sigma_y^2/(\lambda^2(\sigma_{\mu y} - \sigma_y^2)^2)) \quad \text{if } \Delta_{gd} > 0 \quad \text{and } d = 0,$$

$$N^{\Delta_{gd}}(\widehat{\rho}_{AB} - \rho) \xrightarrow{d} N(0, 4/(\lambda\hat{q}(\lambda, c^*))) \quad \text{if } \Delta_{gd} > 0 \quad \text{and } d > 0,$$

$$\widehat{\rho}_{AB} \xrightarrow{d} \rho + \frac{\hat{X}_1}{\hat{X}_2}, \quad \text{with } \begin{bmatrix}\hat{X}_1 \\ \hat{X}_2\end{bmatrix} \sim N\left[\begin{pmatrix} 0 \\ \lambda(\sigma_{\mu y} - \sigma_y^2)/\sigma^2 \end{pmatrix}, (\sigma_y^2/\sigma^2)\begin{pmatrix} 2 & -1 \\ -1 & 1 \end{pmatrix}\right] \quad \text{if } \Delta_{gd} = 0 \text{ and } d = 0,$$

$$\widehat{\rho}_{AB} \xrightarrow{d} \rho + \frac{\hat{X}_1}{\hat{X}_2}, \quad \text{with } \begin{bmatrix}\hat{X}_1 \\ \hat{X}_2\end{bmatrix} \sim N\left[-\frac{\hat{q}(\lambda,c^*)}{2}\begin{pmatrix}0\\1\end{pmatrix}, \frac{\hat{q}(\lambda,c^*)}{2\lambda}\begin{pmatrix}2 & -1\\ -1 & 1\end{pmatrix}\right]$$

https://doi.org/10.1017/S0266466608090531 Published online by Cambridge University Press

$$\text{if } \Delta_{gd} = 0 \quad \text{and} \quad d > 0,$$
$$\widehat{\rho}_{AB} \xrightarrow{d} \rho + \frac{\hat{X}_1}{\hat{X}_2}, \quad \text{with} \begin{bmatrix} \hat{X}_1 \\ \hat{X}_2 \end{bmatrix} \sim N\left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \frac{1}{2} \begin{pmatrix} 2 & -1 \\ -1 & 1 \end{pmatrix} \right] \quad \text{if } \Delta_{gd} < 0.$$

$$(ii)\ \sqrt{N}(\widehat{\rho}_{Arbov} - \rho) \xrightarrow{d} \frac{\hat{X}_3}{\hat{X}_4}, \quad \text{with} \begin{bmatrix} \hat{X}_3 \\ \hat{X}_4 \end{bmatrix}$$
$$\sim N\left[\frac{1}{2} \begin{pmatrix} 0 \\ [2 - q(\lambda, c^+)] \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & \frac{q(\lambda,c)}{2\lambda} \end{pmatrix} \right].$$

Note that $\Delta_{gd} = 1/2(1+d) - g$ and that $\Delta_{gd} < 1/2$. When the values of d and g are such that $\Delta_{gd} = 0$, ρ is weakly identified by the AB moment condition(s). When $\Delta_{gd} > 0$, ρ is nearly weakly identified, and when $\Delta_{gd} < 0$, ρ is nearly nonidentified by the AB moment condition(s).

When $0 \leq d \leq 1$ and $\Delta_{gd} = 0$, there are two extreme cases: if $d = 0$, then one obtains the local-to-nonidentification asymptotic distribution of $\widehat{\rho}_{AB}$ for $g = 1/2$. This case corresponds to local-to-zero asymptotics; cf. Staiger and Stock (1997). On the other hand, if $d = 1$, then one obtains the local-to-nonidentification asymptotic distribution of $\widehat{\rho}_{AB}$ for $g = 1$. When $\Delta_{gd} = 0$ and the value of d increases from zero to one, both the signal, $\mathrm{E}(y_{i,1}\Delta y_{i,2}) = O(N^{d-g})$, and the noise of the AB moment function, $(\mathrm{Var}(y_{i,1}\Delta y_{i,2}))^{1/2} = O(N^{(1/2)d})$, become stronger.

Some of the diagonal path local (-to-nonidentification) asymptotic results of Theorem 4 are very similar to the sequential local asymptotic results of Theorem 1, which were obtained for $g = 1$. If $d = 1$ and $\Delta_{gd} = 0$ (so that $g = 1$), then $\hat{q}(\lambda, c^*) = q(\lambda, c)$ and the sequential local (-to-nonidentification) asymptotic distributions of $\widehat{\rho}_{AB}$ and $\widehat{\rho}_{Arbov}$ as $(S, N \to \infty)_{seq}$ could be obtained from Theorem 4 by letting $c \to \infty$ so that $\hat{q}(\lambda, c^*) = q(\lambda, c) = q(\lambda, c^+) = 1$. Furthermore, when $d > 1$ diagonal path local (-to-nonidentification) asymptotic distributions of $\widehat{\rho}_{AB}$ and $\widehat{\rho}_{Arbov}$ are still obtained for $g = 1$, and they can also be shown to be equal to the sequential local (-to-nonidentification) asymptotic distributions of $\widehat{\rho}_{AB}$ and $\widehat{\rho}_{Arbov}$ as $(S, N \to \infty)_{seq}$ for $g = 1$. On the other hand, the first-order large N fixed S asymptotic distribution of $\widehat{\rho}_{Arbov}$ for $\rho = 1$ can be obtained from Theorem 4 by assuming that $c^+ = 0$, i.e., $d < g \leq 1$, so that $q(\lambda, c^+) = q(\lambda, c) = 0$. Finally, note that $\partial \hat{q}(\lambda, c^*)/(\partial c^*) > 0$.

If $\Delta_{gd} = 0$ and $d = 0$, the limit density of $\widehat{\rho}_{AB}$ given in Theorem 4 is bimodal if and only if $\lambda|\sigma_{\mu y} - \sigma_y^2|/(\sigma_y \sigma) > 2$, whereas if $\Delta_{gd} = 0$ and $d > 0$, the limit density of $\widehat{\rho}_{AB}$ is bimodal if and only if $\sqrt{\lambda \hat{q}(\lambda, c^*)/2} > 2$. In both cases the second mode, if it exists, is not of practical importance. When $\Delta_{gd} < 0$, the limit density of $\widehat{\rho}_{AB}$ is unimodal. The limit density of $\widehat{\rho}_{Arbov}$ is always unimodal.

To see that the preceding diagonal path local asymptotic results can yield reasonable approximations, consider the following example: let $\rho = 0.95$, $N = 100$, and $S = 10$. Choose $d = g = 1$. Then $\lambda = 5$, $c = 1/10$, and $q(\lambda, c)N/(2\lambda) = (1 - \exp(-1)) \times 10 = 6.32$, whereas $(1 - \rho^{2(1+S)})/(1 - \rho^2) = 6.93$.

https://doi.org/10.1017/S0266466608090531 Published online by Cambridge University Press

In the remainder of this section it is assumed that $\operatorname{plim}_{N\to\infty,S/N^d\to c^*} W_N = W$ and $\sqrt{N}(\widehat{\rho}_1-\rho) \xrightarrow{d} \mathcal{K}\neq 0$ as $N,S\to\infty$ with $S/N^d\to c^*$, and $\operatorname{plim}_{N\to\infty}$ is short for $\operatorname{plim}_{N\to\infty,S/N^d\to c^*}$. The following lemma and two theorems extend some of the diagonal path local (-to-nonidentification) asymptotic results of Theorem 4 to $T\geq 3$ and to the System estimator.

LEMMA 2. *Let $T\geq 3$, $0\leq d\leq 1$, and $g=1/2(1+d)$. Furthermore, let $\rho = 1-\lambda N^{-g}$ with $\lambda>0$. Let $N\to\infty$, $S\to\infty$ simultaneously with $S/N\to c\geq 0$ and $S/N^d\to c^*>0$. Let $\hat{q}(\lambda,c^*)=2\lambda(S+1)$ if $d=0$, $\hat{q}(\lambda,c^*)=2\lambda c^*$ if $0<d<1$, and $\hat{q}(\lambda,c^*)=q(\lambda,c^*)$ if $d=1$. Finally, let $\sigma_d^4=\sigma^2(\sigma^2\hat{q}(\lambda,c^*)/(2\lambda)+1\{d=0\}\sigma^2_{-S})$. Then*

(i) (1) $N^{(-1/2)(1+d)}\sum_{i=1}^N Z_i^{I\prime}\Delta v_i \xrightarrow{d} \hat{X}_{51}$,
$N^{(-1/2)(1+d)}\sum_{i=1}^N Z_i^{I\prime}\Delta y_{i,-1} \xrightarrow{d} \hat{X}_{61}$,
$N^{-1/2}\sum_{i=1}^N Z_i^{D\prime}\Delta v_i \xrightarrow{d} \hat{X}_{52}$, *and*
$N^{-1/2}\sum_{i=1}^N Z_i^{D\prime}\Delta y_{i,-1} \xrightarrow{d} \hat{X}_{62}$ *with*
$\hat{X}_5=(\hat{X}_{51}'\ \hat{X}_{52}')'\sim N(0,\hat{\Sigma}_{55})$, $\hat{X}_6=(\hat{X}_{61}'\ \hat{X}_{62}')'\sim N(\hat{\mu}_6,\hat{\Sigma}_{66})$, *and* $\mathrm{E}(\hat{X}_5\hat{X}_6')=\hat{\Sigma}_{56}\neq 0$, *where* $\hat{\mu}_6$, $\hat{\Sigma}_{55}$, $\hat{\Sigma}_{56}$, *and* $\hat{\Sigma}_{66}$ *are given in the proof;*

(2) when $T=3$, $\hat{X}_5=\hat{X}_{51}=\sigma^2\hat{X}_1$ *and* $\hat{X}_6=\hat{X}_{61}=\sigma^2\hat{X}_2$, *where* $\hat{X}_1$ *and* $\hat{X}_2$ *are defined in Theorem 4.*

(ii) (1) $N^{-1/2}\sum_{i=1}^N Z_i^{L\prime}v_i \xrightarrow{d} \hat{X}_{71}$ *and*
$N^{-1}\sum_{i=1}^N Z_i^{L\prime}y_{i,-1} \xrightarrow{d} \hat{X}_{81}\sim N(\sigma^2(1-q(\lambda,c^+)/2)\iota,$
$(\sigma^4 q(\lambda,c)/(2\lambda))I)$ *with*
$\hat{X}_7=(\hat{X}_{52}'\ \hat{X}_{71}')'\sim N(0,\hat{\Sigma}_{77})$, $\hat{\Sigma}_{77}=\sigma^4 K^{II}K^{II\prime}$, *and* $\mathrm{E}(\hat{X}_7\hat{X}_8')=0$, *where* $\hat{X}_8=(0'\ \hat{X}_{81}')'$;

(2) when $T=3$, $\hat{X}_7=\hat{X}_{71}=\sigma^2\hat{X}_3$ *and* $\hat{X}_8=\hat{X}_{81}=\sigma^2\hat{X}_4$, *where* $\hat{X}_3$ *and* $\hat{X}_4$ *are defined in Theorem 4.*

In addition, we have the following results concerning the weight matrices:

(iii) $\operatorname{plim}_{N\to\infty}N^{-(1+d)}\sum_{i=1}^N(Z_i^{I\prime}Z_i^I)=(\sigma_d^4/\sigma^2)I$,
$\operatorname{plim}_{N\to\infty}N^{(-1/2)(2+d)}\sum_{i=1}^N(Z_i^{I\prime}Z_i^D)=0$, *and*
$\operatorname{plim}_{N\to\infty}N^{-1}\sum_{i=1}^N(Z_i^{D\prime}Z_i^D)=\sigma^2 I$.

(iv) $\operatorname{plim}_{N\to\infty}W_{N,Arbov1}=\sigma^{-2}I$ *and*

$$N^{-1}\sum_{i=1}^N\left(\tilde{Z}_i^{II\prime}[\widehat{\Delta v_i'}\ \widehat{v_i'}]'[\widehat{\Delta v_i'}\ \widehat{v_i'}]\tilde{Z}_i^{II}\right)\xrightarrow{d}\sigma^4\hat{W}_{Arbov2}^{-1}$$

$$\equiv\sigma^4K^{II}K^{II\prime}+(\sigma^4/(2\lambda))q(\lambda,c)\mathcal{K}^2\operatorname{diag}(O_{m-(T-2)},I_{T-2}).$$

Proof. See the complete version of the paper available at http://webspace.qmul.ac.uk/hkruiniger/. ■

https://doi.org/10.1017/S0266466608090531 Published online by Cambridge University Press

THEOREM 5. *Let $T \geq 3$, $0 \leq d \leq 1$, and $g = 1/2(1+d)$. Furthermore, let $\rho = 1 - \lambda N^{-g}$ with $\lambda > 0$. Let $N \to \infty$, $S \to \infty$ simultaneously with $S/N \to c \geq 0$ and $S/N^d \to c^* > 0$. Then*

(i) (1) $\widehat{\rho}_{AB1} - \rho \xrightarrow{d} \hat{X}_6' \hat{\Sigma}_{55}^{-1} \hat{X}_5 / \hat{X}_6' \hat{\Sigma}_{55}^{-1} \hat{X}_6$. *If $T > 3$, $\widehat{\rho}_{AB1}$ is asymptotically biased downward.*

(2) assuming that $\widehat{\rho}_{AB}$ exploits $\mathrm{E}(Z_i^{AB\prime} \Delta v_i) = 0$ in lieu of $\mathrm{E}(Z_i' \Delta v_i) = 0$,

$$\widehat{\rho}_{AB} - \rho \xrightarrow{d} \hat{X}_{14}' W \hat{X}_{13} / \hat{X}_{14}' W \hat{X}_{14}$$

with $\hat{X}_{13} = (\hat{X}_{51}' \; 1\{d=0\}\hat{X}_{52}')'$ and $\hat{X}_{14} = (\hat{X}_{61}' \; 1\{d=0\}\hat{X}_{62}')'$.

(ii) $\operatorname{plim}_{N\to\infty}(\widehat{\rho}_{Arbov} - \rho) = 0$, *and assuming that $\widehat{\rho}_{Arbov}$ exploits $\mathrm{E}(\tilde{Z}_i^{II\prime} \times [\Delta v_i' \; v_i']') = 0$ in lieu of $\mathrm{E}(Z_i^{II\prime} v_i) = 0$,*

$$\sqrt{N}(\widehat{\rho}_{Arbov} - \rho) \xrightarrow{d} \hat{X}_8' W \hat{X}_7 / \hat{X}_8' W \hat{X}_8.$$

Proof. See the complete version of the paper available at http://webspace.qmul.ac.uk/hkruiniger/.

When $0 \leq d \leq 1$ and $g = 1/2(1+d)$, both $\widehat{\rho}_{AB1}$ and $\widehat{\rho}_{AB2}$ are inconsistent. Moreover, $\operatorname{plim}_{N\to\infty, S/N^d \to c^*}(\widehat{\rho}_{AB2} - \widehat{\rho}_{AB1}) \neq 0$ because $\widehat{\Delta v_i}(\widehat{\rho}_{AB1}) = \Delta v_i - (\widehat{\rho}_{AB1} - \rho)\Delta y_{i,-1}$ and $\widehat{\rho}_{AB1} - \rho \xrightarrow{d} \hat{X}_6' \hat{\Sigma}_{55}^{-1} \hat{X}_5 / \hat{X}_6' \hat{\Sigma}_{55}^{-1} \hat{X}_6 \neq 0$. ■

THEOREM 6. *Let $T \geq 3$, $0 \leq d \leq 1$, and $g = 1/2(1+d)$. Furthermore, let $\rho = 1 - \lambda N^{-g}$ with $\lambda > 0$. Let $N \to \infty$, $S \to \infty$ simultaneously with $S/N \to c \geq 0$ and $S/N^d \to c^* > 0$. Let $\hat{q}(\lambda, c^*) = 2\lambda(S+1)$ if $d = 0$, $\hat{q}(\lambda, c^*) = 2\lambda c^*$ if $0 < d < 1$, and $\hat{q}(\lambda, c^*) = q(\lambda, c^*)$ if $d = 1$. Let $\hat{W}_{SYS*} = \operatorname{diag}(1\{d=0\}\sigma_y^{-2} H_{T-2}^{-1}, \sigma^{-2} I_m)$. Finally, let $D = \mathrm{E}[\operatorname{diag}(\varepsilon_i^2, \ldots, \varepsilon_i^{T-1})' \iota \Delta \varepsilon_i'] / \sigma^2$. Then*

(i) $\operatorname{plim}_{N\to\infty} W_{N,SYS1} = \hat{W}_{SYS*}$ *and* $\operatorname{plim}_{N\to\infty} W_{N,SYS1b} = \hat{W}_{SYS*}$.

(ii) $\operatorname{diag}(N^{(-1/2)(1+d)} \; I_{T-2}, \; N^{-1/2} I_m) \; \sum_{i=1}^N (\tilde{Z}_i^{S\prime} \; [\widehat{\Delta v_i'} \; \widehat{v_i'}]' \; [\widehat{\Delta v_i'} \; \widehat{v_i'}] \; \tilde{Z}_i^S)$ $\operatorname{diag}(N^{(-1/2)(1+d)} I_{T-2}, N^{-1/2} I_m) \xrightarrow{d} \sigma^4 \hat{W}_{SYS2}^{-1}$ *with* $\sigma^4 \hat{W}_{SYS2}^{11} \equiv \sigma^2(\sigma^2 \times \hat{q}(\lambda, c^*)/(2\lambda) + 1\{d=0\}\sigma_{-S}^2) H$, $\hat{W}_{SYS2}^{21} = (\hat{W}_{SYS2}^{12})' \equiv -(1/(2\lambda)) q(\lambda, c) \times \mathcal{K} K^{II} D$, *and* $\hat{W}_{SYS2}^{22} \equiv \hat{W}_{Arbov2}^{-1}$, *where $\hat{W}_{Arbov2}^{-1}$ is defined in Lemma 2.*

(iii) $\operatorname{plim}_{N\to\infty}(\widehat{\rho}_{SYSk} - \rho) = 0$, *for $k = 1, 1b, 2$, and*

$$\sqrt{N}(\widehat{\rho}_{SYSk} - \rho) \xrightarrow{d} \hat{X}_8' \hat{\Sigma}_{77}^{-1} \hat{X}_7 / \hat{X}_8' \hat{\Sigma}_{77}^{-1} \hat{X}_8, \quad \text{for } k = 1, 1b,$$
$$\sqrt{N}(\widehat{\rho}_{SYS2} - \rho) \xrightarrow{d} \hat{X}_{11}' \hat{W}_{SYS2} \hat{X}_{10} / \hat{X}_{11}' \hat{W}_{SYS2} \hat{X}_{11}$$

with $\hat{X}_{10} = (\hat{X}_{51}' \; \hat{X}_7')'$ and $\hat{X}_{11} = (0 \; \hat{X}_8')'$, where $\hat{\Sigma}_{77}$, $\hat{X}_{51}$, $\hat{X}_7$, and $\hat{X}_8$ are defined in Lemma 2.

https://doi.org/10.1017/S0266466608090531 Published online by Cambridge University Press

(iv) assuming that $\widehat{\rho}_{SYS}$ *exploits* $\mathrm{E}(\tilde{Z}_i^{S\prime}[\Delta v_i' \;\; v_i']') = 0$ *in lieu of* $\mathrm{E}(Z_i^{S\prime} \times [\Delta v_i' \, v_i']') = 0$,

$$N^{1/2(1-d)}(\widehat{\rho}_{SYS} - \rho) \xrightarrow{d} \hat{X}_{16}' W \hat{X}_{15} / \hat{X}_{16}' W \hat{X}_{16}$$

with $\hat{X}_{15} = (\hat{X}_{51}' \; 1\{d = 0\}\hat{X}_7')'$ *and* $\hat{X}_{16} = (1\{d = 1\}\hat{X}_{61}' \; \hat{X}_8')'$, *where* $\hat{X}_{51}$, $\hat{X}_{61}$, $\hat{X}_7$, *and* $\hat{X}_8$ *are defined in Lemma 2.*

Proof. See the complete version of the paper available at http://webspace.qmul.ac.uk/hkruiniger/. ■

When $d > 1$ the diagonal path local (-to-nonidentification) asymptotic properties of the AB, Arbov, and System estimators mentioned in Theorems 5 and 6 are equal to their sequential local asymptotic properties given in Theorems 2 and 3.

Note that when $d \geq 1$ and $g = 1$ the estimators of the optimal weight matrices for the Arbov estimator and the System estimator and the asymptotic standard errors of the latter estimators are inconsistent and the local asymptotic distribution of the two-step "optimal" System estimator is skewed.

We now consider the diagonal path asymptotic distributions of the AB and Arbov estimators for $\rho = 1$ and $T = 3$.

THEOREM 7. *Let* $T = 3$ *and let* $\rho = 1$. *Furthermore, let* $N \to \infty$, $S \to \infty$ *simultaneously with* $S/N^d \to c^* > 0$ *and* $S/N^{\bar{d}} \to c \geq 0$ *where* $d \geq 0$ *and* $\bar{d} = \max(d, 1)$. *Finally, let* $1\{d \leq 1\} = 1$ *if* $d \leq 1$ *and let* $1\{d \leq 1\} = 0$ *if* $d > 1$. *Then*

(i) $\widehat{\rho}_{AB} - \rho \xrightarrow{d} \frac{\hat{X}_1}{\hat{X}_2}$, *where* $\begin{bmatrix} \hat{X}_1 \\ \hat{X}_2 \end{bmatrix} \sim N\left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 & -1 \\ -1 & 1 \end{pmatrix}\right]$.

(ii) $N^{(1/2)\bar{d}}(\widehat{\rho}_{Arbov} - \rho) \xrightarrow{d} \frac{\hat{X}_3}{\hat{X}_4}$, *where* $\begin{bmatrix} \hat{X}_3 \\ \hat{X}_4 \end{bmatrix} \sim N\left[\begin{pmatrix} 0 \\ 1\{d \leq 1\} \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & c \end{pmatrix}\right]$.

When $\rho = 1$ and $T = 3$, the diagonal path asymptotic distribution of the AB estimator is equal to both its first-order large N fixed S asymptotic distribution and its sequential as $(S, N \to \infty)_{seq}$ asymptotic distribution (cf. (15)), regardless of the value of $d \geq 0$. Furthermore, for $d < 1$, the Arbov estimator has a normal asymptotic distribution that is equal to its first-order large N fixed S asymptotic distribution. However, for $d \geq 1$, the Arbov estimator has a nonnormal asymptotic distribution. Moreover, for $d > 1$, the diagonal path asymptotic distribution of the Arbov estimator is equal to its sequential as $(S, N \to \infty)_{seq}$ asymptotic distribution, which is a Cauchy distribution; cf. (16).

The $\sqrt{S}$ rate (the faster than $\sqrt{N}$ rate) of convergence and the nonnormal shape of the limiting distribution that are obtained for $\widehat{\rho}_{Arbov}$ when $\rho = 1$ and $S, N \to \infty$ with $S/N^d \to c^* > 0$ and $d > 1$ or $(S, N \to \infty)_{seq}$ are determined/caused

https://doi.org/10.1017/S0266466608090531 Published online by Cambridge University Press

by the behavior of the sum that involves the initial conditions, which dominates the sum(s) that only involve(s) the sample period innovations; see the proof of Theorem 7 or note 5. Also when $d = 1$, the asymptotic distribution of $\widehat{\rho}_{Arbov}$ is affected by the behavior of the initial conditions. Phillips and Magdalinos (2008) observe similar "tail wagging the dog" phenomena in the asymptotic behavior of the LS estimator for a simple unit root AR model with a so-called infinite past initialization and the related t-ratio. Note that although the initial conditions also play a dominating role in the derivation of the asymptotic distribution of $\widehat{\rho}_{AB}$ when $S, N \to \infty$ with $S/N^d \to c^* > 0$ and $d > 1$ or $(S, N \to \infty)_{seq}$, the distribution obtained is the same as the one found for any d with $0 \leq d < 1$.

Observe that $\lim_{\lambda \downarrow 0} \hat{q}(\lambda, c^*) = 0$ and $\lim_{\lambda \downarrow 0}(q(\lambda, c)/2\lambda) = \lim_{\lambda \downarrow 0}((1 - \exp(-2\lambda c))/(2\lambda)) = \lim_{\lambda \downarrow 0}(2c \exp(-2\lambda c)/2) = c$. From these observations and the results in Theorems 4 and 7 we conclude that the diagonal path local asymptotic distributions of the Arbov and System estimators are left-continuous at $\lambda = 0$ ($\rho = 1$) provided that $0 \leq d \leq 1$. A similar continuity result holds for the local-to-nonidentification asymptotic distribution of the AB estimator for any $d \geq 0$.

The limit densities of $\widehat{\rho}_{AB}$ and $\widehat{\rho}_{Arbov}$ in Theorem 7 are both unimodal.

The following lemma and theorem extend the results in Theorem 7 to $T > 3$ and the System estimator.

LEMMA 3. *Let $T \geq 3$ and let $\rho = 1$. Furthermore, let $N \to \infty$, $S \to \infty$ simultaneously with $S/N^d \to c^* > 0$ and $S/N^{\bar{d}} \to c \geq 0$ where $d \geq 0$ and $\bar{d} = \max(d, 1)$. Let $\sigma_{ur}^4 = \sigma^2(c^*\sigma^2 + 1\{d = 0\}(\sigma_{-S}^2 + \sigma^2))$. Finally, let $\widehat{\rho}_{1b}$ be an initial $N^{(1/2)\bar{d}}$-consistent estimator for ρ, i.e., $N^{(1/2)\bar{d}}(\widehat{\rho}_{1b} - \rho) \xrightarrow{d} \mathcal{K}_b \neq 0$. Then*

(i) (1) $N^{(-1/2)(1+d)} \sum_{i=1}^N Z_i^{I\prime} \Delta v_i \xrightarrow{d} \hat{X}_{51}$,
$N^{(-1/2)(1+d)} \sum_{i=1}^N Z_i^{I\prime} \Delta y_{i,-1} \xrightarrow{d} \hat{X}_{61}$,
$N^{-1/2} \sum_{i=1}^N Z_i^{D\prime} \Delta v_i \xrightarrow{d} \hat{X}_{52}$, *and*
$N^{-1/2} \sum_{i=1}^N Z_i^{D\prime} \Delta y_{i,-1} \xrightarrow{d} \hat{X}_{62}$ *with*
$\hat{X}_5 = (\hat{X}_{51}' \ \hat{X}_{52}')' \sim N(0, \hat{\Sigma}_{55})$, $\hat{X}_6 = (\hat{X}_{61}' \ \hat{X}_{62}')' \sim N(\hat{\mu}_6, \hat{\Sigma}_{66})$, *and* $\mathrm{E}(\hat{X}_5 \hat{X}_6') = \hat{\Sigma}_{56} \neq 0$, *where* $\hat{\mu}_6$, $\hat{\Sigma}_{55}$, $\hat{\Sigma}_{56}$, *and* $\hat{\Sigma}_{66}$ *are given in the proof;*

(2) when $T = 3$, $\hat{X}_5 = \hat{X}_{51} = \sigma_{ur}^2 \hat{X}_1$ *and* $\hat{X}_6 = \hat{X}_{61} = \sigma_{ur}^2 \hat{X}_2$, *where* $\hat{X}_1$ *and* $\hat{X}_2$ *are defined in Theorem 7.*

(ii) (1) $N^{-1/2} \sum_{i=1}^N Z_i^{L\prime} v_i \xrightarrow{d} \hat{X}_{71}$ *and*
$N^{(-1/2)(1+\bar{d})} \sum_{i=1}^N Z_i^{L\prime} y_{i,-1} \xrightarrow{d} \hat{X}_{81} \sim N(1\{d \leq 1\}\sigma^2 \iota, c\sigma^4 I)$ *with*
$\hat{X}_7 = (\hat{X}_{52}' \ \hat{X}_{71}')' \sim N(0, \hat{\Sigma}_{77})$ *and* $\mathrm{E}(\hat{X}_7 \hat{X}_8') = 0$, *where* $\hat{\Sigma}_{77}$ *is defined in Lemma 2 and* $\hat{X}_8 = (0' \ \hat{X}_{81}')'$;

(2) when $T = 3$, $\hat{X}_7 = \hat{X}_{71} = \sigma^2 \hat{X}_3$, *and* $\hat{X}_8 = \hat{X}_{81} = \sigma^2 \hat{X}_4$, *where* $\hat{X}_3$ *and* $\hat{X}_4$ *are defined in Theorem 7.*

https://doi.org/10.1017/S0266466608090531 Published online by Cambridge University Press

In addition, we have the following results concerning the weight matrices:

(iii) $\operatorname{diag}(N^{-1/2}I_{m-(T-2)}, N^{(-1/2)\bar{d}}I_{T-2})\sum_{i=1}^{N}(\tilde{Z}_i^{II\prime}[\widehat{\Delta v_i'}(\widehat{\rho}_1)\,\widehat{v}_i'(\widehat{\rho}_1)]'$
$\times[\widehat{\Delta v_i'}(\widehat{\rho}_1)\,\widehat{v}_i'(\widehat{\rho}_1)]\tilde{Z}_i^{II})\operatorname{diag}(N^{-1/2}I_{m-(T-2)}, N^{(-1/2)\bar{d}}I_{T-2}) \xrightarrow{d} \sigma^4\hat{W}_{Arbov2}^{-1}$,
$\operatorname{diag}(N^{-1/2}I_{m-(T-2)}, N^{-1/2}I_{T-2})\sum_{i=1}^{N}(\tilde{Z}_i^{II\prime}[\widehat{\Delta v_i'}(\widehat{\rho}_{1b})\,\widehat{v}_i'(\widehat{\rho}_{1b})]'[\widehat{\Delta v_i'}(\widehat{\rho}_{1b})$
$\widehat{v}_i'(\widehat{\rho}_{1b})]\tilde{Z}_i^{II})\operatorname{diag}(N^{-1/2}I_{m-(T-2)}, N^{-1/2}I_{T-2}) \xrightarrow{d} \sigma^4\hat{W}_{Arbov2b}^{-1}$,
$\operatorname{diag}(N^{(-1/2)(1+d)}I_{T-2}, N^{-1/2}I_{m-(T-2)}, N^{(-1/2)\bar{d}}I_{T-2})$
$\times\sum_{i=1}^{N}(\tilde{Z}_i^{S\prime}[\widehat{\Delta v_i'}(\widehat{\rho}_1)\,\widehat{v}_i'(\widehat{\rho}_1)]'[\widehat{\Delta v_i'}(\widehat{\rho}_1)\,\widehat{v}_i'(\widehat{\rho}_1)]\tilde{Z}_i^{S})$
$\times\operatorname{diag}(N^{(-1/2)(1+d)}I_{T-2}, N^{-1/2}I_{m-(T-2)}, N^{(-1/2)\bar{d}}I_{T-2}) \xrightarrow{d} \sigma^4\hat{W}_{SYS2}^{-1}$,
$\operatorname{diag}(N^{(-1/2)(1+d)}I_{T-2},\ N^{-1/2}I_{m-(T-2)}, N^{-1/2}I_{T-2})$
$\times\sum_{i=1}^{N}(\tilde{Z}_i^{S\prime}[\widehat{\Delta v_i'}(\widehat{\rho}_{1b})\,\widehat{v}_i'(\widehat{\rho}_{1b})]'\ [\widehat{\Delta v_i'}(\widehat{\rho}_{1b})\,\widehat{v}_i'(\widehat{\rho}_{1b})]\tilde{Z}_i^{S})$
$\times\operatorname{diag}(N^{(-1/2)(1+d)}I_{T-2},\ N^{-1/2}I_{m-(T-2)}, N^{-1/2}I_{T-2}) \xrightarrow{d} \sigma^4\hat{W}_{SYS2b}^{-1}$,
where $\hat{W}_{Arbov2}^{-1}$, $\hat{W}_{Arbov2b}^{-1}$, $\hat{W}_{SYS2}^{-1}$, *and* $\hat{W}_{SYS2b}^{-1}$ *are PD matrices.*

THEOREM 8. *Let* $T \geq 3$ *and let* $\rho = 1$. *Furthermore, let* $N \to \infty$, $S \to \infty$ *simultaneously with* $S/N^d \to c^* > 0$ *and* $S/N^{\bar{d}} \to c \geq 0$ *where* $d \geq 0$ *and* $\bar{d} = \max(d, 1)$. *Finally, let* $\widehat{\rho}_{1b}$ *be an initial* $N^{(1/2)\bar{d}}$*-consistent estimator for* ρ, *i.e.,* $N^{(1/2)\bar{d}}(\widehat{\rho}_{1b} - \rho) \xrightarrow{d} \mathcal{K}_b \neq 0$. *Then*

(i) *(1)* $\widehat{\rho}_{AB1} - \rho \xrightarrow{d} \hat{X}_6'\hat{\Sigma}_{55}^{-1}\hat{X}_5/\hat{X}_6'\hat{\Sigma}_{55}^{-1}\hat{X}_6$.
(2) assuming that $\widehat{\rho}_{AB}$ *exploits* $\mathrm{E}(Z_i^{AB\prime}\Delta v_i) = 0$ *in lieu of* $\mathrm{E}(Z_i'\Delta v_i) = 0$,

$$\widehat{\rho}_{AB} - \rho \xrightarrow{d} \hat{X}_{14}'W\hat{X}_{13}/\hat{X}_{14}'W\hat{X}_{14}$$

with $\hat{X}_{13} = (\hat{X}_{51}'\ 1\{d=0\}\hat{X}_{52}')'$ *and* $\hat{X}_{14} = (\hat{X}_{61}'\ 1\{d=0\}\hat{X}_{62}')'$.

(ii) $\operatorname{plim}_{N\to\infty}(\widehat{\rho}_{Arbov} - \rho) = 0$, *and assuming that* $\widehat{\rho}_{Arbov}$ *exploits* $\mathrm{E}(\tilde{Z}_i^{II\prime}$ $\times[\Delta v_i'\ v_i']') = 0$ *in lieu of* $\mathrm{E}(Z_i^{II\prime}v_i) = 0$,

$$N^{(1/2)\bar{d}}(\widehat{\rho}_{Arbov} - \rho) \xrightarrow{d} \hat{X}_8'W\hat{X}_7/\hat{X}_8'W\hat{X}_8.$$

(iii) $\operatorname{plim}_{N\to\infty}(\widehat{\rho}_{SYSk} - \rho) = 0$, *for* $k = 1, 1b, 2, 2b$, *and*

$$N^{(1/2)\bar{d}}(\widehat{\rho}_{SYSk} - \rho) \xrightarrow{d} \hat{X}_8'\hat{\Sigma}_{77}^{-1}\hat{X}_7/\hat{X}_8'\hat{\Sigma}_{77}^{-1}\hat{X}_8, \quad \text{for } k = 1, 1b,$$

$$\sqrt{N}(\widehat{\rho}_{SYS2}(\widehat{\rho}_1) - \rho) \xrightarrow{d} \hat{X}_{11}'\hat{W}_{SYS2}\hat{X}_{10}/\hat{X}_{11}'\hat{W}_{SYS2}\hat{X}_{11},$$

$$N^{(1/2)\bar{d}}(\widehat{\rho}_{SYS2}(\widehat{\rho}_{1b}) - \rho) \xrightarrow{d} \hat{X}_{11}'\hat{W}_{SYS2b}\hat{X}_{10b}/\hat{X}_{11}'\hat{W}_{SYS2b}\hat{X}_{11}$$

with $\hat{X}_{10} = (\hat{X}_{51}'\ \hat{X}_{52}'\ 1\{d \leq 1\}\hat{X}_{71}')'$, $\hat{X}_{10b} = (\hat{X}_{51}'\ \hat{X}_7')'$, *and* $\hat{X}_{11} = (0\ \hat{X}_8')'$.

(iv) *(1) if* $d < 1$, $\operatorname{plim}_{N\to\infty}(\widehat{\rho}_{SYS} - \rho) = 0$.

https://doi.org/10.1017/S0266466608090531 Published online by Cambridge University Press

(2) assuming that $\widehat{\rho}_{SYS}$ *exploits* $\mathrm{E}(\tilde{Z}_i^{S\prime}[\Delta v_i' \; v_i']') = 0$ *in lieu of* $\mathrm{E}(Z_i^{S\prime} \times [\Delta v_i' \; v_i']') = 0$,

$$N^{(1/2)(\bar{d}-d)}(\widehat{\rho}_{SYS} - \rho) \xrightarrow{d} \hat{X}_{16}' W \hat{X}_{15} / \hat{X}_{16}' W \hat{X}_{16}$$

with $\hat{X}_{15} = (\hat{X}_{51}' \; 1\{d = 0\}\hat{X}_7')'$ *and* $\hat{X}_{16} = (1\{d \geq 1\}\hat{X}_{61}' \; \hat{X}_8')'$.

When $\rho = 1$ and $T > 3$, the diagonal path asymptotic distribution of the $\hat{\rho}_{AB1}$ estimator is equal to both its first-order large N fixed S asymptotic distribution and its sequential as $(S, N \to \infty)_{seq}$ asymptotic distribution for all $d \geq 0$. However, in general, these equalities no longer hold simultaneously for any $d \geq 0$ when the AB estimator uses an arbitrary sequence of PD weight matrices. Then one needs to distinguish between two cases: the diagonal path asymptotic distribution of the AB estimator is equal to its first-order large N fixed S asymptotic distribution when $d = 0$, and equal to its sequential as $(S, N \to \infty)_{seq}$ asymptotic distribution when $d > 0$. Finally, when $\rho = 1$ and $T > 3$, the Arbov estimator that uses an arbitrary sequence of PD weight matrices has a normal asymptotic distribution that is equal to its first-order large N fixed S asymptotic distribution for $d < 1$, whereas it has a nonnormal asymptotic distribution for $d \geq 1$, just as in the case $\rho = 1$ and $T = 3$; cf. the discussion following Theorem 7. Moreover, for $d > 1$, the diagonal path asymptotic distribution of the Arbov estimator is equal to its sequential as $(S, N \to \infty)_{seq}$ asymptotic distribution.

4. GMM-BASED PANEL UNIT ROOT TESTS

4.1. LM Tests

In this section we propose two LM-type panel unit root (UR) test statistics that are based on an Arbov estimator and a System estimator that use a weight matrix that is optimal under the null, and on restricted conventional estimators of their first-order fixed-parameter asymptotic standard errors. The critical values for these LM tests can be taken from the standard normal distribution irrespective of the assumptions made regarding the initial observations.

Let $\widehat{\rho}_{Arbov2,R}$ and $\widehat{\rho}_{SYS2,R}$ be GMM estimators that use $W_{N,Arbov2}(1)$ and $W_{N,SYS2}(1)$ as weight matrix, respectively. Let $SE(\widehat{\rho}_{Arbov2,R}) = \{N^{-1}\sum_{i=1}^N (y_{i,-1}' Z_i^{II})[W_{N,Arbov2}(1)]\sum_{i=1}^N (Z_i^{II\prime} y_{i,-1})\}^{-1/2}$ and $SE(\widehat{\rho}_{SYS2,R}) = \{N^{-1}\sum_{i=1}^N [(\Delta y_{i,-1}' \; y_{i,-1}')Z_i^S][W_{N,SYS2}(1)]\sum_{i=1}^N [Z_i^{S\prime}(\Delta y_{i,-1}' y_{i,-1}')']\}^{-1/2}$. Then we have the following results.

THEOREM 9. *Let* $T \geq 3$.

(i) If $\rho = 1$ *and if* $N \to \infty$, $S \to \infty$ *simultaneously with* $S/N^d \to c^* > 0$ *for some* $d \geq 0$, *then*

$$(\widehat{\rho}_{Arbov2,R} - 1)/SE(\widehat{\rho}_{Arbov2,R}) \xrightarrow{d} N(0,1)$$

https://doi.org/10.1017/S0266466608090531 Published online by Cambridge University Press

and

$$(\widehat{\rho}_{SYS2,R}-1)/SE(\widehat{\rho}_{SYS2,R}) \xrightarrow{d} N(0,1).$$

(ii) If $\rho = 1-\lambda N^{-1/2}$ *with* $\lambda > 0$ *and if* $S \to \infty$, $N \to \infty$ *sequentially, then*

$$(\widehat{\rho}_{Arbov2,R}-1)/SE(\widehat{\rho}_{Arbov2,R}) \xrightarrow{d} N\left(-\frac{\lambda}{2}\sqrt{m},1\right)$$

and

$$(\widehat{\rho}_{SYS2,R}-1)/SE(\widehat{\rho}_{SYS2,R}) \xrightarrow{d} N\left(-\frac{\lambda}{2}\sqrt{m},1\right).$$

(iii) If $\rho = 1-\lambda N^{-1/2}$ *with* $\lambda \geq 0$ *and if* $N \to \infty$, $S \to \infty$ *simultaneously with* $S/N^{1/2} \to c \geq 0$ *and* $S/N^{d} \to c^{*} > 0$ *where* $d \geq 0$, *then*

$$(\widehat{\rho}_{Arbov2,R}-1)/SE(\widehat{\rho}_{Arbov2,R}) \xrightarrow{d} N(-\lambda(1-q(\lambda,c)/2)\sqrt{m},1)$$

and

$$(\widehat{\rho}_{SYS2,R}-1)/SE(\widehat{\rho}_{SYS2,R}) \xrightarrow{d} N(-\lambda(1-q(\lambda,c)/2)\sqrt{m},1).$$

The results in Theorem 9(iii) corresponding to $c = 0$ are also valid under large N fixed S asymptotics. Madsen (2003) has derived some related local power results.

Note that the local power is the lowest when the data are CS or $d > 1/2$.

Any Arbov estimator can be used to construct an LM-type panel unit root test, but choosing the optimal Arbov estimator $\widehat{\rho}_{Arbov2,R}$ yields a test statistic with the highest local power within this class of LM tests. In particular, in the proof of Theorem 9 it is shown that the weight matrix used by $\widehat{\rho}_{Arbov2,R}$ is not only optimal under the null but also optimal under local alternatives given by $\rho = 1-\lambda N^{-1/2}$ with $\lambda > 0$, irrespective of the asymptotic plan for S and N. On the other hand, not every System estimator is suitable for the construction of an LM-type panel unit root test that has correct size for any value of d and nontrivial power against any local alternative. However, the optimal System estimator $\widehat{\rho}_{SYS2,R}$ is $\sqrt{N}$-consistent under both the null and local alternatives and yields an LM test that has correct size and the same local power properties as the LM test that is based on the optimal Arbov estimator $\widehat{\rho}_{Arbov2,R}$, irrespective of the asymptotic plan for S and N.

In the last decade various other panel unit root tests have been proposed. For instance, Breitung and Meyer (1994) proposed a test statistic that is based on an ordinary least squares estimator for ρ in a model for deviations from the initial observations. Harris and Tzavalis (1999) discussed an LM-type panel unit root test that is based on the bias-corrected least squares dummy variables (LSDV) estimator for ρ. Finally, Kruiniger (2008b) discussed a Wald-type panel unit root test that is based on the first difference maximum likelihood estimator (FDMLE) for ρ in the covariance stationary panel AR(1)/unit root model.

https://doi.org/10.1017/S0266466608090531 Published online by Cambridge University Press

4.2. Monte Carlo Results

In this section we compare the finite-sample performance of our GMM-based panel unit root tests with three other panel unit root tests, namely, the Wald test that is based on the FDMLE, the LM test due to Harris and Tzavalis (1999) that is based on the bias-corrected LSDV estimator, and a Wald test that is based on the fixed effects maximum likelihood estimator (FEMLE) for ρ in the panel AR(1) model (cf. Hsiao et al., 2002; Kruiniger, 2003, 2008b).[8] We only consider size-adjusted versions of the latter two tests.[9] The implementation of the FEMLE-based unit root test is further discussed in Kruiniger (2008b).

In most simulation experiments the errors have been drawn from normal distributions: $\varepsilon_{i,t} \sim N(0,1)$ and $\mu_i \sim N(0,1)$. To assess how assumptions with respect to $y_{i,1} - \mu_i, i = 1, \ldots, N$, affect the power of the tests, we have conducted four different kinds of experiments: in one set the initial observations are drawn from stationary distributions, i.e., $(y_{i,1} - \mu_i)|\mu_i \sim N(0, 1/(1-\rho^2))$, whereas in the other three sets the initial observations are nonstationary. The three nonstationary cases considered are (1) $y_{i,1} - \mu_i = 0$, (2) $(y_{i,1} - \mu_i)|\mu_i \sim N(0, 2/(1-\rho^2))$, and (3) $(y_{i,1} - 2\mu_i)|\mu_i \sim N(0, \rho^2/(1-\rho^2))$. Note that in all situations $\mathrm{E}(y_{i,t} - y_{i,t-1}) = 0$ as is the case under the null hypothesis. In both case (1) and case (2) the variance of $y_{i,1} - \mu_i$ is different from the variance under stationarity, whereas in case (3) nonstationarity is due to the fact that $\mathrm{E}[\mu_i(y_{i,1} - \mu_i)] \neq 0$. Case (1) corresponds to small S. We have also considered experiments with $\sigma_\mu^2 = 0$, $\sigma_\mu^2 = 100$, and $\varepsilon_{i,t} \sim (\chi^2(1) - 1)/\sqrt{2}$. Note that the size and the power of test statistics that only exploit data in differences are not affected by changes in σ_μ^2. Finally, in the simulation experiments we have varied the dimensions of the panel data sets also: $(N, T) = (100, 10)$, $(100, 6)$, or $(500, 6)$. All simulation results are based on 5,000 replications, and the (nominal) level of the tests is either 2.5% or 5%.

Tables 1 and 2 report the simulation results on the empirical size and power of the panel unit root tests that were mentioned previously. Table 1 reports results on power against stationary alternatives, whereas Table 2 reports results on power against nonstationary alternatives. In the tables "W" denotes the Wald version of a test, "LM" stands for the LM version of a test, and "SA" indicates that the test has been size adjusted. When the (nominal) level of a test is 2.5% this is indicated by ^{a}; otherwise the level of a test is 5%. Inspection of the results in Tables 1 and 2 leads to the following conclusions with respect to the GMM-based unit root tests:

1. In most cases considered the GMM-based tests have correct size. However nonnormality of the errors affects the size of the tests.
2. The power of the test based on the System estimator is greater than or equal to the power of the test based on the Arbov estimator. However, in many cases the power of both tests is roughly the same and equal to the power of the FDMLE.
3. When the variance of the $y_{i,1} - \mu_i$ is larger than the value implied by covariance stationarity, the power of the test based on the System estimator

https://doi.org/10.1017/S0266466608090531 Published online by Cambridge University Press

TABLE 1. Size and power against stationary alternatives

T	N	test	size, $S=-1$	size, $S=49$	$\rho=0.95$	$\rho=0.9$	$\rho=0.8$
A. $\sigma^2=1$ and $\sigma_\mu^2=0$							
6	100	ARBOV2-LM[a]	0.035	0.030	0.155	0.355	0.846
		ARBOV2-LM	0.058	0.056	0.228	0.470	0.906
		SYS2-LM[a]	0.029	0.030	0.143	0.370	0.881
		SYS2-LM	0.050	0.054	0.215	0.477	0.930
B. $\sigma^2=1$ and $\sigma_\mu^2=1$							
10	100	ARBOV2-LM[a]	0.030	0.029	0.261	0.663	0.987
		ARBOV2-LM	0.059	0.054	0.370	0.756	0.993
		SYS2-LM[a]	0.024	0.028	0.209	0.592	0.981
		SYS2-LM	0.049	0.047	0.306	0.700	0.989
6	100	ARBOV2-LM[a]	0.035	0.030	0.145	0.339	0.844
		ARBOV2-LM	0.058	0.056	0.217	0.459	0.901
		SYS2-LM[a]	0.029	0.030	0.139	0.341	0.856
		SYS2-LM	0.050	0.054	0.209	0.453	0.916
		FDML-W	0.056	0.056	0.200	0.466	0.925
		FEML-W-SA	0.049	0.049	0.118	0.162	0.319
		LSDV-LM-SA	0.050	0.050	0.153	0.327	0.812
6	500	ARBOV2-LM[a]	0.029	0.029	0.446	0.939	1.000
		ARBOV2-LM	0.055	0.053	0.553	0.968	1.000
		SYS2-LM[a]	0.027	0.028	0.446	0.958	1.000
		SYS2-LM	0.050	0.052	0.567	0.979	1.000
C. $\sigma^2=1$ and $\sigma_\mu^2=100$							
6	100	ARBOV2-LM[a]	0.035	0.030	0.085	0.161	0.376
		ARBOV2-LM	0.058	0.056	0.138	0.236	0.459
		SYS2-LM[a]	0.029	0.030	0.077	0.147	0.351
		SYS2-LM	0.050	0.054	0.127	0.215	0.435
D. $(y_{i,1}-\mu_i)\|\mu_i \sim (x^2(1)-1)\sqrt{2(1-\rho^2)}$ and $\sigma_\mu^2=1$							
6	100	ARBOV2-LM[a]	0.052	0.047	0.163	0.311	0.619
		ARBOV2-LM	0.100	0.091	0.256	0.420	0.732
		SYS2-LM[a]	0.036	0.039	0.141	0.328	0.743
		SYS2-LM	0.070	0.075	0.228	0.447	0.831

is greater than the power of the tests based on the Arbov estimator and the FDMLE.

4. The power of the GMM-based tests decreases with an increase of the value of σ_μ^2.
5. In the cases considered the GMM-based tests have greater power than either the test that is based on the LSDV estimator or the test that is based on the FEMLE.

https://doi.org/10.1017/S0266466608090531 Published online by Cambridge University Press

TABLE 2. Size and power against nonstationary alternatives

T	N	test	size, $S=-1$	size, $S=49$	$\rho=0.95$	$\rho=0.9$	$\rho=0.8$
A. $y_{i,1}=\mu_i$, $\sigma^2=1$, and $\sigma_\mu^2=1$							
6	100	ARBOV2-LM[a]	0.035	0.030	0.308	0.799	0.999
		ARBOV2-LM	0.058	0.056	0.416	0.872	0.999
		SYS2-LM[a]	0.029	0.030	0.281	0.759	0.998
		SYS2-LM	0.050	0.054	0.378	0.842	0.999
		FDML-W	0.056	0.056	0.457	0.914	1.000
		FEML-W-SA	0.049	0.049	0.138	0.182	0.290
		LSDV-LM-SA	0.050	0.050	0.308	0.732	0.994
B. $(y_{i,1}-\mu_i)\vert\mu_i \sim N(0,2/(1-\rho^2))$, $\sigma^2=1$, and $\sigma_\mu^2=1$							
6	100	ARBOV2-LM[a]	0.035	0.030	0.096	0.138	0.359
		ARBOV2-LM	0.058	0.056	0.157	0.217	0.473
		SYS2-LM[a]	0.029	0.030	0.089	0.189	0.609
		SYS2-LM	0.050	0.054	0.151	0.278	0.721
		FDML-W	0.056	0.056	0.068	0.116	0.351
		FEML-W-SA	0.049	0.049	0.098	0.170	0.396
		LSDV-LM-SA	0.050	0.050	0.060	0.101	0.378
C. $(y_{i,1}-2\mu_i)\vert\mu_i \sim N(0,\rho^2/(1-\rho^2))$, $\sigma^2=1$, and $\sigma_\mu^2=1$							
6	100	ARBOV2-LM[a]	0.035	0.030	0.138	0.351	0.799
		ARBOV2-LM	0.058	0.056	0.216	0.459	0.870
		SYS2-LM[a]	0.029	0.030	0.133	0.355	0.826
		SYS2-LM	0.050	0.054	0.205	0.461	0.893
		FDML-W	0.056	0.056	0.205	0.474	0.927
		FEML-W-SA	0.049	0.049	0.118	0.166	0.325
		LSDV-LM-SA	0.050	0.050	0.143	0.335	0.811

5. CONCLUDING REMARKS

In this paper we considered GMM-based estimation and inference for the panel AR(1) model when the data are persistent and the time dimension of the panel is fixed. We derived local asymptotic approximations to the finite-sample distributions of the AB, Arbov, and System estimators, respectively, under a variety of distributional assumptions about the initial observations. Among other things we found that the nature of the weak instruments problem of the AB estimator depends on the distributional properties of the initial observations. Moreover, when $\rho=1-\lambda/N$ and when either the data are covariance stationary or both S and N grow large with $S/N^d \to c^* > 0$ and $d \geq 1$, then both the Arbov and the two-step "optimal" System estimator have nonnormal local asymptotic distributions, and the estimators of the optimal weight matrices for the Arbov estimator and the System estimator and their asymptotic standard errors are no longer consistent. We also argued that in these cases one should use LM tests and not Wald tests.

https://doi.org/10.1017/S0266466608090531 Published online by Cambridge University Press

Two LM-type panel unit root tests that we proposed were found to have good size and power properties.

NOTES

1. We assume identical distributions across the individuals for convenience.

2. Note that $\mathrm{E}(\varepsilon_{i,t} y_{i,s}) = 0$ for all t and all $s < t$ and $\mathrm{E}(\varepsilon_{i,t}\varepsilon_{i,s}) = 0$ for all $t \neq s$ implies $\mathrm{E}(\varepsilon_{i,t}\eta_i) = 0$ for all $t \geq -S+2$ but not $\mathrm{E}(\varepsilon_{i,-S+1}\eta_i) = 0$.

3. In Section 3.2 we discuss a range of alternative local asymptotic distributions based on $S/N^d \to c^* > 0$ and $\rho = 1 - \lambda N^{-g}$ with $0 \leq d \leq g \leq 1$ so that $\mathrm{Var}(y_{i,1}) = O(N^d)$; cf. Lemma A.4 in the Appendix. The value of d determines the order of magnitude of $\mathrm{Var}(y_{i,1})$ in terms of N. One can expect that if S or $\mathrm{Var}(y_{i,1})$ is small relative to N, then a local asymptotics based on $S/N^d \to c^*$ with d close to zero gives a better approximation to the distributions of the AB and Arbov estimators compared to local asymptotics based on the CS model, which implies that $\mathrm{Var}(y_{i,1})$ is large relative to N.

4. Note that $\mathrm{E}(dm_{AB,s,t}/d\rho|\rho = 1) = \mathrm{E}(y_{i,s}\varepsilon_{i,t-1}) = 0$ when $s \leq t-2$ and that $\mathrm{E}(dm_{Arbov,s,t}/d\rho|\rho = 1) = \mathrm{E}(y_{i,t-1}\varepsilon_{i,s}) = \sigma^2$ when $s \leq t-1$. This implies that $\mathrm{E}(dm_{AB,s,t}/d\rho)$ and $\mathrm{E}(dm_{Arbov,s,t}/d\rho)$ are left-discontinuous at $\rho = 1$.

5. When $\rho = 1$ and $S \to \infty$, the $y_{i,t}$ are no longer tight random variables. However, the results in (15) and (16) can still be shown to be correct upon scaling both the numerator and the denominator in $\widehat{\rho}_{AB} - \rho = (\sum_{i=1}^N y_{i,1}\Delta\varepsilon_{i,3})/(\sum_{i=1}^N y_{i,1}\varepsilon_{i,2})$ and the denominator in $\widehat{\rho}_{Arbov} - \rho = (\sum_{i=1}^N \varepsilon_{i,3}\varepsilon_{i,2})/(\sum_{i=1}^N y_{i,2}\varepsilon_{i,2})$ by $\sqrt{S}$.

6. $\mathrm{E}(1/\tilde{X}_8' \tilde{\Sigma}_{77}^{-1} \tilde{X}_8) \leq \mathrm{E}(\tilde{X}_8' W \tilde{\Sigma}_{77} W \tilde{X}_8/(\tilde{X}_8' W \tilde{X}_8)^2)$ for any W.

7. Conditional on $\tilde{X}_8$, these LM test statistics have a standard normal asymptotic distribution. As the latter distribution does not depend on $\tilde{X}_8$, it follows that the unconditional asymptotic distribution of these LM test statistics is also equal to the standard normal distribution.

8. Bond, Nauges, and Windmeijer (2005) have also compared various panel unit root tests by calculating their asymptotic local power and by conducting simulation experiments. The only GMM-based test they consider is a Wald test based on the "optimal" System estimator. As expected this test has poor size and power properties. They also find that the Breitung–Meyer test and the test based on the FDMLE have very similar size and power properties.

9. Note that the FEMLE is different from the LSDV estimator.

REFERENCES

Ahn, S.C. & P. Schmidt (1995) Efficient estimation of models for dynamic panel data. *Journal of Econometrics* 68, 5–28.

Ahn, S.C. & P. Schmidt (1997) Efficient estimation of dynamic panel data models: Alternative assumptions and simplified estimation. *Journal of Econometrics* 76, 309–321.

Arellano, M. & 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.

Arellano, M. & O. Bover (1995) Another look at the instrumental variable estimation of error-components models. *Journal of Econometrics* 68, 29–51.

Blundell, R.W. & S. Bond (1998) Initial conditions and moment restrictions in dynamic panel data models. *Journal of Econometrics* 87, 115–143.

Bond, S.R., C. Nauges, & F. Windmeijer (2005) Unit Roots: Identification and Testing in Micro Panels. CEMMAP Working paper 07/05, Institute for Fiscal Studies, London.

Breitung, J. & W. Meyer (1994) Testing for unit roots using panel data: Are wages on different bargaining levels cointegrated? *Applied Economics* 26, 353–361.

Fieller, E.C. (1932) The distribution of an index in a normal bivariate population. *Biometrika* 24, 428–440.

https://doi.org/10.1017/S0266466608090531 Published online by Cambridge University Press

Forchini, G. (2006) On the bimodality of the exact distribution of the TSLS estimator. *Econometric Theory* 22, 932–946.
Hahn, J., J. Hausman, & G. Kuersteiner (2001) Bias Corrected Instrumental Variables Estimation for Dynamic Panel Models with Fixed Effects. Working paper, M.I.T.
Han, C. & P.C.B. Phillips (2006) GMM with many moment conditions. *Econometrica* 74, 147–192.
Han, C. & P.C.B. Phillips (2007) GMM Estimation for Dynamic Panels with Fixed Effects and Strong Instruments at Unity. Cowles Foundation Discussion paper 1599, Yale University.
Harris, R.D.F. & E. Tzavalis (1999) Inference for unit roots in dynamic panels where the time dimension is fixed. *Journal of Econometrics* 91, 201–226.
Hillier, G.H. (2006) Yet more on the exact properties of IV estimators. *Econometric Theory* 22, 913–931.
Hsiao, C., M.H. Pesaran, & A.K. Tahmiscioglu (2002) Maximum likelihood estimation of fixed effects dynamic panel models covering short time periods. *Journal of Econometrics* 109, 107–150.
Kruiniger, H. (2003) On the Estimation of Panel Regression Models with Fixed Effects. Revised version of Working paper 450, Queen Mary, University of London.
Kruiniger, H. (2008a) GMM Estimation and Inference in Dynamic Panel Data Models with Persistent Data. Revised version of Working paper 428, Queen Mary, University of London. Available at http://webspace.qmul.ac.uk/hkruiniger/.
Kruiniger, H. (2008b) Maximum likelihood estimation and inference methods for the covariance stationary panel AR(1)/unit root model. *Journal of Econometrics* 144, 447–464.
Madsen, E. (2003) Using GMM when Testing for a Unit Root in Panels Where the Time-Dimension is Fixed. CAM Working paper 2003-11, University of Copenhagen.
Marsaglia, G. (1965) Ratios of normal variables and ratios of sums of uniform variables. *Journal of the American Statistical Association* 60, 193–204.
Moon, H.R. & P.C.B. Phillips (2000) Estimation of autoregressive roots near unity using panel data. *Econometric Theory* 16, 927–997.
Phillips, P.C.B. (2006) A remark on bimodality and weak instrumentation in structural equation estimation. *Econometric Theory* 22, 947–960.
Phillips, P.C.B. & C. Han (2008) Gaussian inference in AR(1) time series with or without a unit root. *Econometric Theory* 24, 631–650.
Phillips, P.C.B. & T. Magdalinos (2008) Unit Root and Cointegrating Limit Theory when Initialization is in the Infinite Past. Cowles Foundation Discussion paper 1655, Yale University.
Phillips, P.C.B. & H.R. Moon (1999) Linear regression limit theory for nonstationary panel data. *Econometrica* 67, 1057–1111.
Staiger, D. & J.H. Stock (1997) Instrumental variables regression with weak instruments. *Econometrica* 65, 557–586.
Stock, J.H, J.H. Wright, & M. Yogo (2002) A survey of weak instruments and weak identification in generalized method of moments. *Journal of Business & Economic Statistics* 20, 518–529.
Woglom, G. (2001) More results on the exact small sample properties of the instrumental variable estimator. *Econometrica* 69, 1381–1389.

APPENDIX: Proofs of the Results

LEMMA A.1. *Let* $\{y_{i,t}\}$ *be a stationary process and let* $u_{i,t} = y_{i,t} - \mu_i$. *Then*

$$\mathrm{E}(u_{i,t-1}\,\Delta u_{i,t}) = -\frac{\sigma^2}{1+\rho}, \qquad \mathrm{E}[(\Delta u_{i,t})^2] = \frac{2\sigma^2}{1+\rho},$$

$$\mathrm{E}(u_{i,t}^4) = \frac{(1-\rho^2)\kappa + 6\rho^2}{(1-\rho^2)^2(1+\rho^2)}\sigma^4,$$

https://doi.org/10.1017/S0266466608090531 Published online by Cambridge University Press

$$\mathrm{E}[(\Delta u_{i,t})^4] = \frac{(1-\rho)^2[\kappa(1-\rho^2)+6\rho^2]}{(1+\rho)^2(1+\rho^2)}\sigma^4 + \frac{6(1-\rho)}{1+\rho}\sigma^4 + \kappa\sigma^4,$$

$$\mathrm{E}[u_{i,t-1}(\Delta u_{i,t})^3] = -\frac{(1-\rho)[\kappa(1-\rho^2)+6\rho^2]}{(1+\rho)^2(1+\rho^2)}\sigma^4 - \frac{3}{(1+\rho)}\sigma^4, \quad \textit{and}$$

$$\mathrm{E}[u_{i,t-1}^2(\Delta u_{i,t})^2] = \frac{(\kappa(1-\rho^2)+6\rho^2)}{(1+\rho)^2(1+\rho^2)}\sigma^4 + \frac{1}{1-\rho^2}\sigma^4.$$

Proof of Lemma A.1. Note that $u_{i,t} = \rho u_{i,t-1} + \varepsilon_{i,t}$ and $\Delta u_{i,t} = (\rho-1)u_{i,t-1} + \varepsilon_{i,t}$. Moreover $\{u_{i,t}\}$ is a stationary process. Then verification of the first two claims is straightforward. The other claims are proved as follows:

$$\mathrm{E}(u_{i,t}^4) = \mathrm{E}(\rho^4 u_{i,t-1}^4 + 6\rho^2 u_{i,t-1}^2\varepsilon_{i,t}^2 + \varepsilon_{i,t}^4) = \rho^4\mathrm{E}(u_{i,t}^4) + \frac{6\sigma^4\rho^2}{(1-\rho^2)} + \kappa\sigma^4$$

$$\Leftrightarrow \mathrm{E}(u_{i,t}^4) = \frac{\kappa\sigma^4}{1-\rho^4} + \frac{6\sigma^4\rho^2}{(1-\rho^2)(1-\rho^4)} = \frac{(1-\rho^2)\kappa+6\rho^2}{(1-\rho^2)^2(1+\rho^2)}\sigma^4,$$

$$\begin{aligned}\mathrm{E}[(\Delta u_{i,t})^4] &= \mathrm{E}[(\rho-1)^4 u_{i,t-1}^4 + 6(\rho-1)^2 u_{i,t-1}^2\varepsilon_{i,t}^2 + \varepsilon_{i,t}^4] \\ &= \frac{(1-\rho)^2[\kappa(1-\rho^2)+6\rho^2]}{(1+\rho)^2(1+\rho^2)}\sigma^4 + \frac{6(1-\rho)}{1+\rho}\sigma^4 + \kappa\sigma^4,\end{aligned}$$

$$\begin{aligned}\mathrm{E}[u_{i,t-1}(\Delta u_{i,t})^3] &= \mathrm{E}[(\rho-1)^3 u_{i,t-1}^4 + 3(\rho-1)u_{i,t-1}^2\varepsilon_{i,t}^2] \\ &= -\frac{(1-\rho)[\kappa(1-\rho^2)+6\rho^2]}{(1+\rho)^2(1+\rho^2)}\sigma^4 - \frac{3}{(1+\rho)}\sigma^4, \quad \text{and}\end{aligned}$$

$$\mathrm{E}[u_{i,t-1}^2(\Delta u_{i,t})^2] = \mathrm{E}[(\rho-1)^2 u_{i,t-1}^4 + u_{i,t-1}^2\varepsilon_{i,t}^2] = \frac{(\kappa(1-\rho^2)+6\rho^2)}{(1+\rho)^2(1+\rho^2)}\sigma^4 + \frac{1}{1-\rho^2}\sigma^4.$$

■

LEMMA A.2. *Let $\{y_{i,t}\}$ be a stationary process. Then*

$$\mathrm{E}(y_{i,1}\Delta y_{i,2}) = -\frac{\sigma^2}{1+\rho}, \quad \mathrm{E}(y_{i,2}\Delta y_{i,2}) = \frac{\sigma^2}{1+\rho},$$

$$\mathrm{Var}(y_{i,1}\Delta\varepsilon_{i,3}) = \frac{2\sigma^2[\sigma^2+\sigma_\mu^2(1-\rho^2)]}{(1+\rho)(1-\rho)},$$

$$\mathrm{Var}[(\varepsilon_{i,3}+(1-\rho)\mu_i)\Delta y_{i,2}] = \frac{2\sigma^2[\sigma^2+\sigma_\mu^2(1-\rho)^2]}{1+\rho},$$

$$\begin{aligned}\mathrm{Var}(y_{i,1}\Delta y_{i,2}) = {} & \frac{(\kappa(1-\rho^2)+6\rho^2)\sigma^4}{(1+\rho)^2(1+\rho^2)} + \frac{\sigma^4}{(1+\rho)(1-\rho)} \\ & + \frac{\sigma^2[2\sigma_\mu^2(1+\rho)-\sigma^2]}{(1+\rho)^2},\end{aligned}$$

https://doi.org/10.1017/S0266466608090531 Published online by Cambridge University Press

$$\mathrm{Var}(y_{i,2}\Delta y_{i,2}) = \frac{\rho^2(\kappa(1-\rho^2)+6\rho^2)}{(1+\rho)^2(1+\rho^2)}\sigma^4 + \frac{\sigma^4}{(1+\rho)(1-\rho)} + \frac{\kappa(1+\rho)^2 - 1 - 6\rho(1+\rho) + 2(\sigma_\mu^2/\sigma^2)(1+\rho)}{(1+\rho)^2}\sigma^4,$$

$$\mathrm{Cov}(y_{i,1}\Delta\varepsilon_{i,3}, y_{i,1}\Delta y_{i,2}) = -\frac{\sigma^2[\sigma^2+\sigma_\mu^2(1-\rho^2)]}{(1-\rho)(1+\rho)}, \quad \textit{and}$$

$$\mathrm{Cov}[(\varepsilon_{i,3}+(1-\rho)\mu_i)\Delta y_{i,2}, y_{i,2}\Delta y_{i,2}] = \frac{2\sigma^2\sigma_\mu^2(1-\rho)}{1+\rho}.$$

Proof of Lemma A.2. Noting that $y_{i,t} = u_{i,t} + \mu_i$, application of Lemma A.1 yields

$$\mathrm{E}(y_{i,1}\Delta y_{i,2}) = \mathrm{E}[(u_{i,1}+\mu_i)\Delta u_{i,2}] = -\frac{\sigma^2}{1+\rho},$$

$$\mathrm{E}(y_{i,2}\Delta y_{i,2}) = \mathrm{E}[y_{i,1}\Delta y_{i,2} + (\Delta u_{i,2})^2] = \frac{\sigma^2}{1+\rho},$$

$$\mathrm{Var}(y_{i,1}\Delta\varepsilon_{i,3}) = \mathrm{E}[(y_{i,1}\Delta\varepsilon_{i,3})^2] = \mathrm{E}[((u_{i,1}+\mu_i)\Delta\varepsilon_{i,3})^2] = \frac{2\sigma^4}{1-\rho^2} + 2\sigma_\mu^2\sigma^2 = \frac{2\sigma^2[\sigma^2+\sigma_\mu^2(1-\rho^2)]}{(1+\rho)(1-\rho)},$$

$$\mathrm{Var}[(\varepsilon_{i,3}+(1-\rho)\mu_i)\Delta y_{i,2}] = \mathrm{E}[((\varepsilon_{i,3}+(1-\rho)\mu_i)\Delta u_{i,2})^2] = \frac{2\sigma^2[\sigma^2+\sigma_\mu^2(1-\rho)^2]}{1+\rho},$$

$$\mathrm{Var}(y_{i,1}\Delta y_{i,2}) = \mathrm{E}\left[(u_{i,1}\Delta u_{i,2} + \mu_i\Delta u_{i,2} + \frac{\sigma^2}{1+\rho})^2\right] = \mathrm{E}\left[(u_{i,1}\Delta u_{i,2})^2 - \frac{2\sigma^4}{(1+\rho)^2} + \frac{2\sigma^2\sigma_\mu^2}{1+\rho} + \frac{\sigma^4}{(1+\rho)^2}\right] = \frac{(\kappa(1-\rho^2)+6\rho^2)\sigma^4}{(1+\rho)^2(1+\rho^2)} + \frac{\sigma^4}{(1+\rho)(1-\rho)} + \frac{\sigma^2[2\sigma_\mu^2(1+\rho)-\sigma^2]}{(1+\rho)^2},$$

$$\mathrm{Var}(y_{i,2}\Delta y_{i,2}) = \mathrm{E}\left[(u_{i,1}\Delta u_{i,2} + (\Delta u_{i,2})^2 + \mu_i\Delta u_{i,2} - \frac{\sigma^2}{1+\rho})^2\right] = \mathrm{E}[(u_{i,1}\Delta u_{i,2})^2 + (\Delta u_{i,2})^4 + (\mu_i\Delta u_{i,2})^2 + \frac{\sigma^4}{(1+\rho)^2} + 2u_{i,1}(\Delta u_{i,2})^3$$

https://doi.org/10.1017/S0266466608090531 Published online by Cambridge University Press

$$
\begin{aligned}
&- \frac{2\sigma^2}{1+\rho}(u_{i,1}\Delta u_{i,2}) - \frac{2\sigma^2}{1+\rho}(\Delta u_{i,2})^2] \\
&= \frac{(\kappa(1-\rho^2)+6\rho^2)}{(1+\rho)^2(1+\rho^2)}\sigma^4 \\
&\quad + \frac{\sigma^4}{1-\rho^2} + \frac{(1-\rho)^2[\kappa(1-\rho^2)+6\rho^2]}{(1+\rho)^2(1+\rho^2)}\sigma^4 \\
&\quad + \frac{6(1-\rho)}{1+\rho}\sigma^4 + \kappa\sigma^4 + \frac{2\sigma^2\sigma_\mu^2}{1+\rho} + \frac{\sigma^4}{(1+\rho)^2} \\
&\quad - \frac{2(1-\rho)[\kappa(1-\rho^2)+6\rho^2]}{(1+\rho)^2(1+\rho^2)}\sigma^4 \\
&\quad - \frac{6}{(1+\rho)}\sigma^4 + \frac{2\sigma^4}{(1+\rho)^2} - \frac{4\sigma^4}{(1+\rho)^2} \\
&= \frac{\rho^2(\kappa(1-\rho^2)+6\rho^2)}{(1+\rho)^2(1+\rho^2)}\sigma^4 \\
&\quad + \frac{\sigma^4}{(1+\rho)(1-\rho)} \\
&\quad + \frac{\kappa(1+\rho)^2 - 1 - 6\rho(1+\rho) + 2(\sigma_\mu^2/\sigma^2)(1+\rho)}{(1+\rho)^2}\sigma^4,
\end{aligned}
$$

$$
\begin{aligned}
\mathrm{Cov}(y_{i,1}\Delta\varepsilon_{i,3}, y_{i,1}\Delta y_{i,2}) &= \mathrm{Cov}(u_{i,1}\Delta\varepsilon_{i,3}, u_{i,1}\Delta u_{i,2}) \\
&\quad + \mathrm{Cov}(\mu_i\Delta\varepsilon_{i,3}, \mu_i\Delta u_{i,2}) \\
&= \mathrm{E}(u_{i,1}^2\Delta\varepsilon_{i,3}\Delta u_{i,2}) + \sigma_\mu^2\mathrm{E}(\Delta\varepsilon_{i,3}\Delta u_{i,2}) \\
&= -\mathrm{E}(u_{i,1}^2\varepsilon_{i,2}^2) - \sigma_\mu^2\mathrm{E}(\varepsilon_{i,2}^2) \\
&= -\frac{\sigma^2[\sigma^2+\sigma_\mu^2(1-\rho^2)]}{(1-\rho)(1+\rho)}, \quad \textit{and}
\end{aligned}
$$

$$
\mathrm{Cov}[(\varepsilon_{i,3}+(1-\rho)\mu_i)\Delta y_{i,2}, y_{i,2}\Delta y_{i,2}] = \mathrm{E}[(1-\rho)(\mu_i\Delta u_{i,2})^2] = \frac{2\sigma^2\sigma_\mu^2(1-\rho)}{1+\rho}. \qquad \blacksquare
$$

A.1. Proof of Theorem 1. When $T=3$, $\widehat{\rho}_{AB} = \rho + (N^{-1}\sum_{i=1}^N y_{i,1}\Delta\varepsilon_{i,3})/(N^{-1}\sum_{i=1}^N y_{i,1}\Delta y_{i,2})$, and $\widehat{\rho}_{Arbov} = \rho + (N^{-1}\sum_{i=1}^N(\varepsilon_{i,3}+(1-\rho)\mu_i)\Delta y_{i,2})/(N^{-1}\sum_{i=1}^N y_{i,2}\Delta y_{i,2})$.

Let us define $X_1 = N^{-1}\sum_{i=1}^N y_{i,1}\Delta\varepsilon_{i,3}$, $X_2 = N^{-1}\sum_{i=1}^N y_{i,1}\Delta y_{i,2}$, $X_3 = N^{-1/2}\sum_{i=1}^N(\varepsilon_{i,3}+(1-\rho)\mu_i)\Delta y_{i,2}$, and $X_4 = N^{-1}\sum_{i=1}^N y_{i,2}\Delta y_{i,2}$.

Let the CS model hold. Then using the results in Lemma A.2 we obtain as $N\to\infty$ for the parameter sequence $\rho = 1-\lambda N^{-1}$ that $X_1 \xrightarrow{d} N(0,\sigma^4/\lambda)$, $X_2 \xrightarrow{d} N(-\sigma^2/2, \sigma^4/(2\lambda))$,

https://doi.org/10.1017/S0266466608090531 Published online by Cambridge University Press

$X_3 \xrightarrow{d} N(0,\sigma^4)$, and $X_4 \xrightarrow{d} N(\sigma^2/2, \sigma^4/(2\lambda))$. It is also easily verified that $\lim_{N\to\infty} \mathrm{Cov}(X_1, X_2) = -\sigma^4/(2\lambda)$ and $\lim_{N\to\infty} \mathrm{Cov}(X_3, X_4) = 0$. It follows that

$$\widehat{\rho}_{AB} \xrightarrow{d} \rho + \frac{\tilde{X}_1}{\tilde{X}_2}, \quad \text{where} \begin{bmatrix} \tilde{X}_1 \\ \tilde{X}_2 \end{bmatrix} \sim N\left[-\frac{\sigma^2}{2}\sqrt{\lambda}\begin{pmatrix} 0 \\ 1 \end{pmatrix}, \frac{\sigma^4}{2}\begin{pmatrix} 2 & -1 \\ -1 & 1 \end{pmatrix} \right], \quad \text{and}$$

$$\sqrt{N}(\widehat{\rho}_{Arbov} - \rho) \xrightarrow{d} \frac{\tilde{X}_3}{\tilde{X}_4}, \quad \text{where} \begin{bmatrix} \tilde{X}_3 \\ \tilde{X}_4 \end{bmatrix} \sim N\left[\frac{\sigma^2}{2}\begin{pmatrix} 0 \\ 1 \end{pmatrix}, \sigma^4\begin{pmatrix} 1 & 0 \\ 0 & \frac{1}{2\lambda} \end{pmatrix} \right]. \quad \blacksquare$$

LEMMA A.3. *Let $\{y_{i,t}\}$ be a stationary process, let $u_{i,t} = y_{i,t} - \mu_i$, let l, p, and q be positive integers, and let $\kappa_q = \left|\mathrm{E}(\varepsilon_{i,t}^q)/\sigma^q\right| < \infty$. Then assuming that the sum of the powers of $\varepsilon_{i,k}$, $u_{i,s}$, $u_{i,t}$, and $\Delta u_{i,t}$ is less than or equal to 8, we have*

(i) $\lim_{\rho\uparrow 1}(1-\rho)^p \mathrm{E}(u_{i,t}^{2p}) < \infty$,
(ii) $\lim_{\rho\uparrow 1}(1-\rho)^p \left|\mathrm{E}(u_{i,t}^{2p+1})\right| < \infty$,
(iii) $\lim_{\rho\uparrow 1} \mathrm{E}[(\Delta u_{i,t})^p] = \kappa_p \sigma^p$,
(iv) $\lim_{\rho\uparrow 1}(1-\rho)^p \left|\mathrm{E}[u_{i,s}^{2p}(\Delta u_{i,t})^q]\right| < \infty$,
(v) $\lim_{\rho\uparrow 1}(1-\rho)^p \left|\mathrm{E}[u_{i,s}^{2p+1}(\Delta u_{i,t})^q]\right| < \infty$,
(vi) $\lim_{\rho\uparrow 1}(1-\rho)^p \left|\mathrm{E}[\varepsilon_{i,k}^l u_{i,s}^{2p}(\Delta u_{i,t})^q]\right| < \infty$, *and*
(vii) $\lim_{\rho\uparrow 1}(1-\rho)^p \left|\mathrm{E}[\varepsilon_{i,k}^l u_{i,s}^{2p+1}(\Delta u_{i,t})^q]\right| < \infty$.

Proof of Lemma A.3. Note that $\{u_{i,t}\}$ is a stationary process, $u_{i,t} = \rho u_{i,t-1} + \varepsilon_{i,t}$, and $\Delta u_{i,t} = (\rho - 1)u_{i,t-1} + \varepsilon_{i,t}$. Moreover, $\mathrm{E}(u_{i,t}^p) = \mathrm{E}[(\rho u_{i,t-1} + \varepsilon_{i,t})^p] = \mathrm{E}[\sum_{k=0}^p B(p,k) \times (\rho u_{i,t-1})^{p-k} \varepsilon_{i,t}^k]$, where $B(p,k) = [k!(p-k)!]^{-1} p!$.

We prove (i) and (ii) together. The proof proceeds by induction:

First consider $p = 0$: $\mathrm{E}(u_{i,t}^0) = 1$ and $\mathrm{E}(u_{i,t}^1) = 0$, and hence $\lim_{\rho\uparrow 1} \mathrm{E}(u_{i,t}^0) < \infty$ and $\lim_{\rho\uparrow 1} \left|\mathrm{E}(u_{i,t}^1)\right| < \infty$.

Now let $p > 0$ and suppose that $\lim_{\rho\uparrow 1}(1-\rho)^q \mathrm{E}(u_{i,t}^{2q}) < \infty$ and $\lim_{\rho\uparrow 1}((1-\rho)^q \times |\mathrm{E}(u_{i,t}^{2q+1})|) < \infty$ for $q \le p-1$. Note that $\lim_{\rho\uparrow 1}(1-\rho^{2p})^{-1}(1-\rho) = \lim_{\rho\uparrow 1}(\sum_{k=0}^{2p-1} \rho^k)^{-1} = 1/(2p)$ and $\lim_{\rho\uparrow 1}(1-\rho^{2p+1})^{-1}(1-\rho) = 1/(2p+1)$. It follows that $\lim_{\rho\uparrow 1}(1-\rho)^p \mathrm{E}(u_{i,t}^{2p}) = \lim_{\rho\uparrow 1}(1-\rho^{2p})^{-1}(1-\rho)^p B(2p,2)\mathrm{E}(u_{i,t}^{2p-2})\mathrm{E}(\varepsilon_{i,t}^2) = \sigma^2(2p)^{-1}B(2p,2) \times \lim_{\rho\uparrow 1}(1-\rho)^{p-1}\mathrm{E}(u_{i,t}^{2p-2}) < \infty$. It also follows that $\lim_{\rho\uparrow 1}(1-\rho)^p \left|\mathrm{E}(u_{i,t}^{2p+1})\right| \le \lim_{\rho\uparrow 1}(1-\rho^{2p+1})^{-1}(1-\rho)^p B(2p+1,2)\left|\mathrm{E}(u_{i,t}^{2p-1})\right|\mathrm{E}(\varepsilon_{i,t}^2) + \lim_{\rho\uparrow 1}(1-\rho^{2p+1})^{-1}(1-\rho)^p B(2p+1,3)\mathrm{E}(u_{i,t}^{2p-2})\left|\mathrm{E}(\varepsilon_{i,t}^3)\right| = \sigma^2(2p+1)^{-1}B(2p+1,2) \times \lim_{\rho\uparrow 1}(1-\rho)^{p-1}\left|\mathrm{E}(u_{i,t}^{2p-1})\right| + |\kappa_3|\sigma^3(2p+1)^{-1}B(2p+1,3) \times \lim_{\rho\uparrow 1}(1-\rho)^{p-1}\mathrm{E}(u_{i,t}^{2p-2}) < \infty$.

The proofs of (iii)–(vii) are now straightforward:

(iii) $\lim_{\rho\uparrow 1} \mathrm{E}[(\Delta u_{i,t})^p] = \lim_{\rho\uparrow 1} \mathrm{E}[((\rho-1)u_{i,t-1} + \varepsilon_{i,t})^p] = \mathrm{E}[(\varepsilon_{i,t})^p] = \kappa_p \sigma^p$,
(iv) $\lim_{\rho\uparrow 1}(1-\rho)^p \left|\mathrm{E}\left[u_{i,s}^{2p}(\Delta u_{i,t})^q\right]\right| = \lim_{\rho\uparrow 1}(1-\rho)^p \left|\mathrm{E}\left[u_{i,s}^{2p}\varepsilon_{i,t}^q\right]\right| < \infty$,

https://doi.org/10.1017/S0266466608090531 Published online by Cambridge University Press

(v) $\lim_{\rho\uparrow 1}(1-\rho)^p\left|\mathrm{E}\left[u_{i,s}^{2p+1}(\Delta u_{i,t})^q\right]\right| = \lim_{\rho\uparrow 1}(1-\rho)^p\left|\mathrm{E}\left[u_{i,s}^{2p+1}(B(q,1)(\rho-1)\times u_{i,t-1}\varepsilon_{i,t}^{q-1}+\varepsilon_{i,t}^q)\right]\right| < \infty$,

(vi) $\lim_{\rho\uparrow 1}(1-\rho)^p\left|\mathrm{E}\left[\varepsilon_{i,k}^l u_{i,s}^{2p}(\Delta u_{i,t})^q\right]\right| = \lim_{\rho\uparrow 1}(1-\rho)^p\left|\mathrm{E}\left[\varepsilon_{i,k}^l u_{i,s}^{2p}\varepsilon_{i,t}^q\right]\right| < \infty$, and finally

(vii) $\lim_{\rho\uparrow 1}(1-\rho)^p\left|\mathrm{E}\left[\varepsilon_{i,k}^l u_{i,s}^{2p+1}(\Delta u_{i,t})^q\right]\right| = \lim_{\rho\uparrow 1}(1-\rho)^p\left|\mathrm{E}\left[\varepsilon_{i,k}^l u_{i,s}^{2p+1}\times(B(q,1)(\rho-1)u_{i,t-1}\varepsilon_{i,t}^{q-1}+\varepsilon_{i,t}^q)\right]\right| < \infty$. ■

A.2. Proofs of Lemma 1 and Theorem 2

Proof of Lemma 1. Part (i): The proof of the first two results is similar to the proof of the first part of Theorem 1:

For $\rho = 1-\lambda N^{-1}$ $\lim_{N\to\infty}\mathrm{E}(y_{i,1}\Delta y_{i,t-1}) = \lim_{N\to\infty}\rho^{t-3}\mathrm{E}(y_{i,1}\Delta y_{i,2}) = -\frac{1}{2}\sigma^2$ and $\lim_{N\to\infty}N^{-1}\mathrm{Var}(y_{i,1}\Delta v_{i,t}) = \lim_{N\to\infty}N^{-1}\mathrm{E}[(y_{i,1}\Delta\varepsilon_{i,t})^2] = \lim_{N\to\infty}(N^{-1}\times\mathrm{E}[(y_{i,1}\Delta\varepsilon_{i,3})^2]) = \lim_{N\to\infty}2\sigma^2N^{-1}\mathrm{E}(y_{i,1}^2) = \lim_{N\to\infty}2\sigma^2N^{-1}\mathrm{E}(u_{i,1}^2) = \sigma^4/\lambda$.

Note that $\mathrm{Var}(y_{i,1}\Delta y_{i,t-1}) = \mathrm{E}[(u_{i,1}\Delta u_{i,t-1}+\mu_i\Delta u_{i,t-1}+\frac{\sigma^2}{1+\rho}\rho^{t-3})^2]$, $u_{i,t} = \rho u_{i,t-1}+\varepsilon_{i,t}$, and $\Delta u_{i,t} = (\rho-1)u_{i,t-1}+\varepsilon_{i,t}$. Then it is easily seen that for $\rho = 1-\lambda N^{-1}$ $\lim_{N\to\infty}N^{-1}\mathrm{Var}(y_{i,1}\Delta y_{i,t-1}) = \lim_{N\to\infty}N^{-1}\mathrm{E}[(u_{i,1}\Delta u_{i,t-1})^2] = \lim_{N\to\infty}N^{-1}\mathrm{E}\{[(\rho-1)u_{i,1}^2+u_{i,1}\varepsilon_{i,t-1}]^2\} = \lim_{N\to\infty}N^{-1}\mathrm{E}[(u_{i,1}\varepsilon_{i,t-1})^2] = \sigma^4/(2\lambda)$; that if $s<t-1$ and $\rho = 1-\lambda N^{-1}$ $\lim_{N\to\infty}N^{-1}\mathrm{Cov}(y_{i,1}\Delta y_{i,s}, y_{i,1}\Delta y_{i,t-1}) = \lim_{N\to\infty}N^{-1}\mathrm{E}(u_{i,1}^2\Delta u_{i,s}\Delta u_{i,t-1}) = \lim_{N\to\infty}N^{-1}\mathrm{E}(u_{i,1}^2\Delta u_{i,s}\Delta u_{i,s+1}) = \lim_{N\to\infty}N^{-1}\mathrm{E}\{u_{i,1}^2[(\rho-1)u_{s-1}+\varepsilon_s][(\rho-1)(\rho u_{s-1}+\varepsilon_s)]\} = 0$; $\lim_{N\to\infty}N^{-1}\mathrm{Cov}(y_{i,1}\Delta v_{i,t}, y_{i,1}\Delta y_{i,t-1}) = \lim_{N\to\infty}N^{-1}\mathrm{E}(y_{i,1}^2\Delta\varepsilon_{i,t}\Delta u_{i,t-1}) = -\sigma^4/(2\lambda)$; and finally that $\lim_{N\to\infty}N^{-1}\mathrm{Cov}(y_{i,1}\Delta v_{i,t}, y_{i,1}\Delta y_{i,t}) = \lim_{N\to\infty}N^{-1}\mathrm{E}(y_{i,1}^2\Delta\varepsilon_{i,t}\Delta u_{i,t}) = \sigma^4/(2\lambda)$.

Thus for $\rho = 1-\lambda N^{-1}$ $\lim_{N\to\infty}N^{-1}\mathrm{E}(Z_i^{I\prime}\Delta v_i\Delta v_i'Z_i^I) = \lim_{N\to\infty}\sigma^2N^{-1}\mathrm{E}(y_{i,1}^2)\times H = \sigma^4/(2\lambda)H$, $\lim_{N\to\infty}N^{-1}\mathrm{E}(Z_i^{I\prime}\Delta y_{i,-1}\Delta y_{i,-1}'Z_i^I) = \lim_{N\to\infty}\sigma^2N^{-1}\mathrm{E}(y_{i,1}^2)I = \sigma^4/(2\lambda)I$, and $\lim_{N\to\infty}N^{-1}\mathrm{E}(Z_i^{I\prime}\Delta y_{i,-1}\Delta v_i'Z_i^I) = -\lim_{N\to\infty}\sigma^2N^{-1}\mathrm{E}(y_{i,1}^2)C = -\sigma^4/(2\lambda)C$, and hence

$$\begin{bmatrix} N^{-1}\sum_{i=1}^N Z_i^{I\prime}\Delta v_i \\ N^{-1}\sum_{i=1}^N Z_i^{I\prime}\Delta y_{i,-1}\end{bmatrix} \xrightarrow{d} \begin{bmatrix}\tilde{X}_{51}\\ \tilde{X}_{61}\end{bmatrix} \sim N\left[\begin{pmatrix}0\\ -\frac{1}{2}\sigma^2\iota\end{pmatrix}, \begin{pmatrix}\frac{\sigma^4}{2\lambda}\end{pmatrix}\begin{pmatrix}H & -C'\\ -C & I\end{pmatrix}\right].$$

Note that $\tilde{\Sigma}_{56,11} = \mathrm{E}(\tilde{X}_{51}\tilde{X}_{61}') = -\sigma^4/(2\lambda)C' \neq 0$.

Consider now the scaled sums $N^{-1/2}\sum_{i=1}^N Z_i^{D\prime}\Delta v_i$ and $N^{-1/2}\sum_{i=1}^N Z_i^{D\prime}\Delta y_{i,-1}$. For $\rho = 1-\lambda N^{-1}$ we obtain $\lim_{N\to\infty}N^{1/2}\mathrm{E}(Z_i^{D\prime}\Delta y_{i,-1}) = \lim_{N\to\infty}N^{1/2}\mathrm{E}(Z_i^{D\prime}\times\Delta u_{i,-1}) = \lim_{N\to\infty}N^{1/2}\mathrm{E}\{Z_i^{D\prime}[(\rho-1)u_{i,-2}+\varepsilon_{i,-1}]\} = 0$. In addition, $\lim_{N\to\infty}\mathrm{E}(Z_i^{D\prime}\Delta v_i\Delta v_i'Z_i^D) = \lim_{N\to\infty}\sigma^2\mathrm{E}(Z_i^{D\prime}HZ_i^D)$, $\lim_{N\to\infty}\mathrm{E}(Z_i^{D\prime}\Delta y_{i,-1}\Delta y_{i,-1}'Z_i^D) = \lim_{N\to\infty}\sigma^2\mathrm{E}(Z_i^{D\prime}Z_i^D)$, and $\lim_{N\to\infty}\mathrm{E}(Z_i^{D\prime}\Delta y_{i,-1}\Delta v_i'Z_i^D) = -\lim_{N\to\infty}\sigma^2\mathrm{E}(Z_i^{D\prime}\times CZ_i^D)$. Hence for $\rho = 1-\lambda N^{-1}$

https://doi.org/10.1017/S0266466608090531 Published online by Cambridge University Press

$$\begin{bmatrix} N^{-1/2}\sum_{i=1}^{N} Z_i^{D\prime}\Delta v_i \\ \\ N^{-1/2}\sum_{i=1}^{N} Z_i^{D\prime}\Delta y_{i,-1} \end{bmatrix} \xrightarrow{d} \begin{bmatrix} \tilde{X}_{52} \\ \\ \tilde{X}_{62} \end{bmatrix}$$

$$\sim N\left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \sigma^4 \begin{pmatrix} \operatorname{diag}(H_{T-3}, H_{T-4}, \ldots, H_1) & -\operatorname{diag}(C_{T-3}, C_{T-4}, \ldots, C_1)' \\ -\operatorname{diag}(C_{T-3}, C_{T-4}, \ldots, C_1) & I \end{pmatrix}\right].$$

Note that $\tilde{\Sigma}_{56,22} = \mathrm{E}(\tilde{X}_{52}\tilde{X}_{62}') = -\sigma^4 \operatorname{diag}(C_{T-3}, C_{T-4}, \ldots, C_1)' \neq 0$.

Next consider the off-diagonal blocks $\tilde{\Sigma}_{55,12}$, $\tilde{\Sigma}_{66,12}$, $\tilde{\Sigma}_{56,12}$, and $\tilde{\Sigma}_{56,21}$. Because $\lim_{N\to\infty} \mathrm{E}(Z_i^{I\prime}\Delta v_i \Delta v_i' Z_i^D) = \lim_{N\to\infty} \sigma^2 \mathrm{E}(Z_i^{I\prime} H Z_i^D) = \lim_{N\to\infty} \sigma^2 HE(y_{i,1} Z_i^D) = \lim_{N\to\infty} \sigma^2 HE(y_{i,1}[d_1 \Delta y_{i,2} \quad d_2 \rho \Delta y_{i,2} \quad \ldots \quad d_{T-3}\rho^{T-4}\Delta y_{i,2}]) = \lim_{N\to\infty} -\frac{1}{2}\sigma^4 \times H[d_1 d_2 \ \ldots \ d_{T-3}]$, we have $\tilde{\Sigma}_{55,12} = \mathrm{E}(\tilde{X}_{51}\tilde{X}_{52}') = \lim_{N\to\infty} \sigma^2 N^{-1/2} \mathrm{E}(Z_i^{I\prime} H Z_i^D) = 0$. Similarly, $\tilde{\Sigma}_{66,12} = \mathrm{E}(\tilde{X}_{61}\tilde{X}_{62}') = \lim_{N\to\infty} \sigma^2 N^{-1/2} \mathrm{E}(Z_i^{I\prime} Z_i^D) = 0$, $\tilde{\Sigma}_{56,12} = \mathrm{E}(\tilde{X}_{51}\tilde{X}_{62}') = -\lim_{N\to\infty} \sigma^2 N^{-1/2} \mathrm{E}(Z_i^{I\prime} C' Z_i^D) = 0$, and $\tilde{\Sigma}_{56,21} = \mathrm{E}(\tilde{X}_{52}\tilde{X}_{61}') = -\lim_{N\to\infty} \sigma^2 N^{-1/2} \mathrm{E}(Z_i^{D\prime} C' Z_i^I) = 0$.

Finally, it is easily seen that if $T = 3$, then $\tilde{X}_5 = \tilde{X}_{51} = \sigma^2 \tilde{X}_1/\sqrt{\lambda}$ and $\tilde{X}_6 = \tilde{X}_{61} = \sigma^2 \tilde{X}_2/\sqrt{\lambda}$, where $\tilde{X}_1$ and $\tilde{X}_2$ are defined in Theorem 1.

Part (ii): The proof of the first two results is similar to the proof of the last part of Theorem 1 and the proof of part (i) of that theorem:

Stationarity implies that for $\rho = 1 - \lambda N^{-1}$ $\lim_{N\to\infty} \mathrm{E}(y_{i,t}\Delta y_{i,t}) = \lim_{N\to\infty} \mathrm{E}(y_{i,2} \times \Delta y_{i,2}) = \sigma^2/2$ and $\lim_{N\to\infty} N^{-1}\mathrm{Var}(y_{i,t}\Delta y_{i,t}) = \lim_{N\to\infty} N^{-1}\mathrm{Var}(y_{i,2}\Delta y_{i,2}) = \sigma^4/(2\lambda)$. Moreover, if $s < t$, then for $\rho = 1 - \lambda N^{-1}$ $\lim_{N\to\infty} N^{-1}\mathrm{E}(y_{i,s}\Delta y_{i,s} y_{i,t}\Delta y_{i,t}) = \lim_{N\to\infty} N^{-1}\mathrm{E}(u_{i,s}\Delta u_{i,s} u_{i,t}\Delta u_{i,t}) = \lim_{N\to\infty} N^{-1}\mathrm{E}(u_{i,s-1}\Delta u_{i,s} u_{i,t-1}\Delta u_{i,t}) = \lim_{N\to\infty} N^{-1}\mathrm{E}(u_{i,s-1}\Delta u_{i,s} u_{i,s-1}\Delta u_{i,s+1}) = \lim_{N\to\infty} N^{-1}\mathrm{E}\{u_{i,s-1}^2[(\rho-1)u_{s-1} + \varepsilon_s] \times[(\rho-1)(\rho u_{s-1} + \varepsilon_s)]\} = 0$.

Thus for $\rho = 1 - \lambda N^{-1}$ $\lim_{N\to\infty} \mathrm{E}(Z_i^{L\prime} v_i v_i' Z_i^L) = \lim_{N\to\infty} \sigma^2 \mathrm{E}(Z_i^{L\prime} Z_i^L) = \sigma^4 I$, $\lim_{N\to\infty} N^{-1}\mathrm{E}(Z_i^{L\prime} y_{i,-1} y_{i,-1}' Z_i^L) = \sigma^4/(2\lambda) I$, and $\lim_{N\to\infty} N^{-1/2}\mathrm{E}(Z_i^{L\prime} y_{i,-1} v_i' \times Z_i^L) = 0$, and hence

$$\begin{bmatrix} N^{-1/2}\sum_{i=1}^{N} Z_i^{L\prime} v_i \\ \\ N^{-1}\sum_{i=1}^{N} Z_i^{L\prime} y_{i,-1} \end{bmatrix} \xrightarrow{d} \begin{bmatrix} \tilde{X}_{71} \\ \\ \tilde{X}_{81} \end{bmatrix} \sim N\left[\begin{pmatrix} 0 \\ \\ \frac{1}{2}\sigma^2 \iota \end{pmatrix}, \left(\frac{\sigma^4}{2\lambda}\right)\begin{pmatrix} 2\lambda I & 0 \\ 0 & I \end{pmatrix}\right].$$

Note that $\mathrm{E}(\tilde{X}_{71}\tilde{X}_{81}') = 0$. Because $\lim_{N\to\infty} N^{-1/2}\mathrm{E}(Z_i^{D\prime}\Delta v_i y_{i,-1}' Z_i^L) = 0$, we also have $\mathrm{E}(\tilde{X}_{52}\tilde{X}_{81}') = 0$. We conclude that $\mathrm{E}(\tilde{X}_7\tilde{X}_8') = 0$. Furthermore it is easily verified that $\tilde{X}_7 \sim N(0, \tilde{\Sigma}_{77})$ with $\tilde{\Sigma}_{77} = \lim_{N\to\infty} K^{II}\mathrm{E}(Z_i^{II\prime} v_i v_i' Z_i^{II}) K^{II\prime} = \sigma^4 K^{II} K^{II\prime}$.

Finally, it is easily seen that if $T = 3$, then $\tilde{X}_7 = \tilde{X}_{71} = \sigma^2 \tilde{X}_3$ and $\tilde{X}_8 = \tilde{X}_{81} = \sigma^2 \tilde{X}_4$, where $\tilde{X}_3$ and $\tilde{X}_4$ are defined in Theorem 1.

Part (iii): Because $\lim_{N\to\infty} N^{-2}\mathrm{Var}(y_{i,1}^2) < \infty$, $\lim_{N\to\infty} N^{-1}\mathrm{Var}(y_{i,1}\Delta y_{i,t}) < \infty$, and $\lim_{N\to\infty} \mathrm{Var}([\Delta y_{i,t}]^2) < \infty$, it is easily seen that for $\rho = 1 - \lambda N^{-1}$, $N^{-2}\sum_{i=1}^{N} (Z_i^{I\prime} Z_i^I) \xrightarrow{q.m.} \lim_{N\to\infty} N^{-1}\mathrm{E}(Z_i^{I\prime} Z_i^I) = (\sigma^2/(2\lambda))I$, $N^{-3/2}\sum_{i=1}^{N} (Z_i^{I\prime} Z_i^D) \xrightarrow{q.m.} \lim_{N\to\infty} N^{-1/2}\mathrm{E}(Z_i^{I\prime} Z_i^D) = 0$, and $N^{-1}\sum_{i=1}^{N} (Z_i^{D\prime} Z_i^D) \xrightarrow{q.m.} \lim_{N\to\infty} \mathrm{E}(Z_i^{D\prime} Z_i^D) = \sigma^2 I$. The results are obtained by noting that convergence in quadratic mean implies convergence in probability.

https://doi.org/10.1017/S0266466608090531 Published online by Cambridge University Press

Part (iv): Because $\lim_{N\to\infty}\mathrm{Var}([\Delta y_{i,t}]^2) < \infty\ \forall t$, it is easily seen that for $\rho = 1-\lambda N^{-1}$, $N^{-1}\sum_{i=1}^N (Z_i^{II\prime}Z_i^{II}) \overset{q.m.}{\to} \lim_{N\to\infty}\mathrm{E}(Z_i^{II\prime}Z_i^{II}) = \sigma^2 I$ and hence $\mathrm{plim}_{N\to\infty}[N^{-1}\sum_{i=1}^N (Z_i^{II\prime}Z_i^{II})] = \sigma^2 I$.

To establish the second claim of part (iv), note that $\widehat{v}_i = y_i - \widehat{\rho}_1 y_{i,-1} = v_i + (\rho - \widehat{\rho}_1) \times y_{i,-1}$ and $\Delta\widehat{v}_i = \Delta y_i - \widehat{\rho}_1 \Delta y_{i,-1} = \Delta v_i + (\rho - \widehat{\rho}_1)\Delta y_{i,-1}$.

Lemma A.3 implies that $\lim_{N\to\infty}\mathrm{E}(\Delta u_{i,t})^8 < \infty$, $\lim_{N\to\infty} N^{-1}\mathrm{E}[u_{i,s}^2(\Delta u_{i,t})^6] < \infty$, $\lim_{N\to\infty} N^{-2}\mathrm{E}[u_{i,s}^4(\Delta u_{i,t})^4] < \infty$, $\lim_{N\to\infty}\mathrm{E}[(\Delta u_{i,t})^4] < \infty$, $\lim_{N\to\infty}\left|\mathrm{E}[u_{i,s} \times (\Delta u_{i,t})^3]\right| < \infty$, and $\lim_{N\to\infty} N^{-1}\mathrm{E}[u_{i,s}^2(\Delta u_{i,t})^2] < \infty\ \forall s,t$. It follows from the first result, our model assumptions, and the Cauchy–Schwarz inequality that $\lim_{N\to\infty}\mathrm{Var}([v_{i,s}\Delta y_{i,t}]^2) < \infty\ \forall s,t$ and hence $N^{-1}\sum_{i=1}^N (\tilde{Z}_i^{II\prime}[\Delta v_i'\ v_i']'[\Delta v_i'\ v_i']\tilde{Z}_i^{II}) \overset{q.m.}{\to} \lim_{N\to\infty} K^{II}\mathrm{E}(Z_i^{II\prime}v_i v_i' Z_i^{II})K^{II\prime} = \sigma^4 K^{II}K^{II\prime}$. Similarly it follows that $N^{-2} \times \sum_{i=1}^N (Z_i^{L\prime}y_{i,-1}y_{i,-1}'Z_i^L) \overset{q.m.}{\to} \lim_{N\to\infty} N^{-1}\mathrm{E}(Z_i^{L\prime}y_{i,-1}y_{i,-1}'Z_i^L) = \sigma^4/(2\lambda)I$, $N^{-2}\sum_{i=1}^N (Z_i^{D\prime}\Delta y_{i,-1}\Delta y_{i,-1}'Z_i^{D\prime}) \overset{q.m.}{\to} 0$, and $N^{-2}\sum_{i=1}^N (Z_i^{L\prime}y_{i,-1}\Delta y_{i,-1}'Z_i^{D\prime}) \overset{q.m.}{\to} 0$.

Lemma A.3 and our model assumptions also imply that $\lim_{N\to\infty}\left|\mathrm{E}[v_{i,k}u_{i,s}(\Delta u_{i,t})^2]\right| < \infty$ and $\lim_{N\to\infty} N^{-1}\mathrm{E}[v_{i,k}^2 u_{i,s}^2(\Delta u_{i,t})^4] < \infty\ \forall k,s,t$. It follows from these results, our model assumptions, and the Cauchy–Schwarz inequality that $N^{-3/2}\sum_{i=1}^N (Z_i^{L\prime}v_i y_{i,-1}'Z_i^L) \overset{q.m.}{\to} 0$, $N^{-3/2}\sum_{i=1}^N (Z_i^{L\prime}v_i\Delta y_{i,-1}'Z_i^D) \overset{q.m.}{\to} 0$, $N^{-3/2}\sum_{i=1}^N (Z_i^{L\prime}y_{i,-1}\Delta v_i'Z_i^D) \overset{q.m.}{\to} 0$, and $N^{-3/2}\sum_{i=1}^N (Z_i^{D\prime}\Delta y_{i,-1}\Delta v_i'Z_i^D) \overset{q.m.}{\to} 0$.

We conclude from the preceding results that $N^{-1}\sum_{i=1}^N \tilde{Z}_i^{II\prime}[\widehat{\Delta v}_i'\ \widehat{v}_i']'[\widehat{\Delta v}_i'\ \widehat{v}_i']\tilde{Z}_i^{II} \overset{d}{\to} \sigma^4 K^{II}K^{II\prime} + \sigma^4/(2\lambda)\mathcal{K}^2 diag(O_{m-(T-2)},\ I_{T-2})$. ■

Proof of Theorem 2. The proof relies on results in (the proof of) Lemma 1.

Part (i): Recall that $Z_i' = K^{AB}Z_i^{AB\prime}$, where $Z_i^{AB} = [Z_i^I\ \ Z_i^D]$ and $rank(K^{AB}) = m$. Therefore $\widehat{\rho}_{AB1}$ is equal to a GMM estimator that exploits $\mathrm{E}(Z_i^{AB\prime}\Delta v_i) = 0$ and uses the weight matrix $(N^{-1}\sum_{i=1}^N Z_i^{AB\prime}HZ_i^{AB})^{-1}$. Consider the local-to-nonidentification limiting behavior of this weight matrix. Because $\lim_{N\to\infty} N^{-2}\mathrm{Var}(y_{i,1}^2) < \infty$, $\lim_{N\to\infty} N^{-1}\mathrm{Var}(y_{i,1}\Delta y_{i,t}) < \infty$, and $\lim_{N\to\infty}\mathrm{Var}([\Delta y_{i,t}]^2) < \infty$, it is easily seen that for $\rho = 1-\lambda N^{-1}$, $\sigma^2 N^{-2}\sum_{i=1}^N (Z_i^{I\prime}HZ_i^I) \overset{q.m.}{\to} \tilde{\Sigma}_{55,11}$ and $\sigma^2 N^{-3/2}\sum_{i=1}^N (Z_i^{I\prime}HZ_i^D) \overset{q.m.}{\to} 0$. Moreover $\sigma^2 N^{-1}\sum_{i=1}^N (Z_i^{D\prime}HZ_i^D) \overset{q.m.}{\to} \tilde{\Sigma}_{55,22}$. Therefore $\mathrm{plim}_{N\to\infty}\sigma^2 N^{-2}\sum_{i=1}^N (Z_i^{I\prime}HZ_i^I) = \tilde{\Sigma}_{55,11}$, $\mathrm{plim}_{N\to\infty}\sigma^2 N^{-1}\sum_{i=1}^N (Z_i^{D\prime}HZ_i^D) = \tilde{\Sigma}_{55,22}$, and $\mathrm{plim}_{N\to\infty}\sigma^2 N^{-3/2}\sum_{i=1}^N (Z_i^{I\prime}HZ_i^D) = \tilde{\Sigma}_{55,12} = 0$. Furthermore $\tilde{\Sigma}_{55}$ is PD. It follows from the preceding results and (the proof of) Lemma 1 that $\widehat{\rho}_{AB1} - \rho \overset{d}{\to} \tilde{X}_6'\tilde{\Sigma}_{55}^{-1}\tilde{X}_5 / \tilde{X}_6'\tilde{\Sigma}_{55}^{-1}\tilde{X}_6$. Assuming that $\widehat{\rho}_{AB}$ exploits $\mathrm{E}(Z_i^{AB\prime}\Delta v_i) = 0$, we also obtain $\widehat{\rho}_{AB} - \rho \overset{d}{\to} \tilde{X}_{61}'W_{11}\tilde{X}_{51} / \tilde{X}_{61}'W_{11}\tilde{X}_{61}$.

Finally we show that $\widehat{\rho}_{AB1}$ is asymptotically biased downward when $T > 3$:

Note that $\widehat{\rho}_{AB1} = \rho + \{\sum_{t=2}^{T-1}[\widetilde{y}_{-1,t-1}'Z_t(Z_t'Z_t)^{-1}Z_t'\widetilde{y}_{-1,t-1}]\}^{-1}\sum_{t=2}^{T-1}[\widetilde{y}_{-1,t-1}'Z_t \times (Z_t'Z_t)^{-1}Z_t'\widetilde{v}_t]$. Because $\sum_{t=2}^{T-1}[\widetilde{y}_{-1,t-1}'Z_t(Z_t'Z_t)^{-1}Z_t'\widetilde{v}_t] = (\sum_{i=1}^N \Delta y_{i,-1}'Z_i^{AB})[\sum_{i=1}^N (Z_i^{AB\prime}HZ_i^{AB})]^{-1}\sum_{i=1}^N Z_i^{AB\prime}\Delta v_i$ and $\sum_{t=2}^{T-1}[\widetilde{y}_{-1,t-1}'Z_t(Z_t'Z_t)^{-1}Z_t'\widetilde{y}_{-1,t-1}] = (\sum_{i=1}^N \Delta y_{i,-1}'Z_i^{AB})[\sum_{i=1}^N (Z_i^{AB\prime}HZ_i^{AB})]^{-1}\sum_{i=1}^N Z_i^{AB\prime}\Delta y_{i,-1}$, it follows from part (i) of the proof of Lemma 1 that both the numerator and denominator of $\widehat{\rho}_{AB1} - \rho$ converge in distribution.

https://doi.org/10.1017/S0266466608090531 Published online by Cambridge University Press

We first show that $\lim_{N\to\infty} \mathrm{E}[\widetilde{y}'_{-1,t-1} Z_t (Z_t'Z_t)^{-1} Z_t' \widetilde{v}_t] < 0$ for $t = 2, \ldots, T-1$. Note that $\lim_{N\to\infty} \mathrm{E}[\widetilde{y}'_{-1,t-1} Z_t (Z_t'Z_t)^{-1} Z_t' \widetilde{v}_t] = \lim_{N\to\infty} \mathrm{E}\{\sum_{i=1}^N [\widetilde{y}_{i,-1,t-1} z'_{i,t} (Z_t'Z_t)^{-1} \times z_{i,t} \widetilde{\varepsilon}_{i,t}]\} = \lim_{N\to\infty} \sum_{i=1}^N [\mathrm{E}(z'_{i,t}(Z_t'Z_t)^{-1} z_{i,t}) \times \mathrm{E}(\widetilde{y}_{i,-1,t-1}\widetilde{\varepsilon}_{i,t})]$.

Now, $\mathrm{E}(\widetilde{y}_{i,-1,t-1}\widetilde{\varepsilon}_{i,t}) = ((T-t)/(T-t+1))\,\mathrm{E}\Big[(y_{i,t-1} - (1/(T-t))\sum_{s=t}^{T-1} y_{i,s})(\varepsilon_{i,t} - (1/(T-t))\sum_{s=t+1}^{T} \varepsilon_{i,s})\Big] = (1/(T-t+1)) \times \mathrm{E}\Big[-\sum_{k=0}^{T-1-t} \rho^k \varepsilon_{i,t}^2 + (1/(T-t))\sum_{s=t+1}^{T-1} \sum_{k=0}^{T-1-s} \rho^k \varepsilon_{i,s}^2\Big] = -\sigma^2 (1/(T-t+1))(1/(T-t))\Big[\sum_{s=t+1}^{T} \sum_{k=T-1-s+1}^{T-1-t} \rho^k\Big]$. It follows that $\lim_{N\to\infty} \mathrm{E}(\widetilde{y}_{i,-1,t-1}\widetilde{\varepsilon}_{i,t}) = \lim_{\rho\uparrow 1} \mathrm{E}(\widetilde{y}_{i,-1,t-1}\widetilde{\varepsilon}_{i,t}) = -\frac{1}{2}\sigma^2$. We also have $\lim_{N\to\infty} \mathrm{E}(\sum_{i=1}^N z'_{i,t}(Z_t'Z_t)^{-1} z_{i,t}) = \mathrm{E}[\mathrm{tr}(I_{t-1})] = t-1 > 0$.

We conclude that $\lim_{N\to\infty} \sum_{t=2}^{T-1} \mathrm{E}(\widetilde{y}'_{-1,t-1} Z_t (Z_t'Z_t)^{-1} Z_t' \widetilde{v}_t) < 0$.

To complete the proof of the asymptotic biasedness of $\widehat{\rho}_{AB1}$, we write the numerator of $\widehat{\rho}_{AB1} - \rho$, i.e., $\sum_{t=2}^{T-1} [(\sum_{i=1}^N \widetilde{y}_{i,-1,t-1} z'_{i,t})(Z_t'Z_t)^{-1} (\sum_{i=1}^N z_{i,t}\widetilde{\varepsilon}_{i,t})]$, as the sum of two terms:

$$\left\{\sum_{t=2}^{T-1}\left[\sum_{i=1}^{N} z'_{i,t}(Z_t'Z_t)^{-1} z_{i,t}\left(\frac{1}{T-t+1}\right)\right.\right.$$
$$\left.\left. \times \left(-\sum_{k=0}^{T-1-t} \rho^k \varepsilon_{i,t}^2 + \frac{1}{T-t}\sum_{s=t+1}^{T-1}\sum_{k=0}^{T-1-s} \rho^k \varepsilon_{i,s}^2\right)\right]\right\} \tag{A.1}$$

$$+\left\{\sum_{t=2}^{T-1}\left[\widetilde{y}'_{-1,t-1} Z_t (Z_t'Z_t)^{-1} Z_t' \widetilde{\varepsilon}_t - \sum_{i=1}^{N} z'_{i,t}(Z_t'Z_t)^{-1} z_{i,t}\left(\frac{1}{T-t+1}\right)\right.\right.$$
$$\left.\left. \times \left(-\sum_{k=0}^{T-1-t} \rho^k \varepsilon_{i,t}^2 + \frac{1}{T-t}\sum_{s=t+1}^{T-1}\sum_{k=0}^{T-1-s} \rho^k \varepsilon_{i,s}^2\right)\right]\right\}. \tag{A.2}$$

The first term, (A.1), converges in probability to $\lim_{N\to\infty} \sum_{t=2}^{T-1} \mathrm{E}[\widetilde{y}'_{-1,t-1} Z_t (Z_t'Z_t)^{-1} \times Z_t' \widetilde{\varepsilon}_t]$. Because the numerator of $\widehat{\rho}_{AB1} - \rho$, $\sum_{t=2}^{T-1} \widetilde{y}'_{-1,t-1} Z_t (Z_t'Z_t)^{-1} Z_t' \widetilde{\varepsilon}_t$, converges in distribution to a random variable with mean $\lim_{N\to\infty} \sum_{t=2}^{T-1} \mathrm{E}[\widetilde{y}'_{-1,t-1} Z_t (Z_t'Z_t)^{-1} Z_t' \widetilde{\varepsilon}_t]$, the second term, (A.2), converges in distribution to a random variable with mean zero. For convenience we will assume that the $\varepsilon_{i,t}$'s are symmetrically distributed around zero. Doing so does not entail a loss of generality because imposing this assumption does not affect the asymptotic distribution of $\widehat{\rho}_{AB1}$. Noting that the second term does not involve higher powers of $\varepsilon_{i,t}$ (higher than one) and using that the $\varepsilon_{i,t}$'s are symmetrically distributed around zero, it follows that the second term is symmetrically distributed around zero and also asymptotically uncorrelated with the denominator of $\widehat{\rho}_{AB1} - \rho$, $\sum_{t=2}^{T-1} [\widetilde{y}'_{-1,t-1} Z_t \times (Z_t'Z_t)^{-1} Z_t' \widetilde{y}_{-1,t-1}]$. We conclude that the expectation of the second term divided by the denominator converges to zero as $N \to \infty$. The ratio of the first term and the denominator converges to a negative constant divided by a positive random variable and gives rise to the negative bias of $\widehat{\rho}_{AB1}$.

Part (ii): It follows in a straightforward manner from the results (in the proof of) Lemma 1 that $\mathrm{plim}_{N\to\infty}(\widehat{\rho}_{Arbov} - \rho) = 0$ and $\sqrt{N}(\widehat{\rho}_{Arbov} - \rho) \xrightarrow{d} \tilde{X}_8' W \tilde{X}_7 / \tilde{X}_8' W \tilde{X}_8$. ■

https://doi.org/10.1017/S0266466608090531 Published online by Cambridge University Press

A.3. Proof of Theorem 3.

Part (i): $W_{N,SYS1} = (N^{-1}\sum_{i=1}^{N} Z_i^{S\prime} A Z_i^S)^{-1}$ where $Z_i^S = \operatorname{diag}(Z_i^I, Z_i^{II})$ and $A = \begin{bmatrix} H & C \\ C' & I_{T-2} \end{bmatrix}$.

Let $M_{11} = N^{-1}\sum_{i=1}^{N}(Z_i^{I\prime} H Z_i^I)$, $M_{12} = M_{21}' = N^{-1}\sum_{i=1}^{N}(Z_i^{I\prime} C Z_i^{II})$, and $M_{22} = N^{-1}\sum_{i=1}^{N}(Z_i^{II\prime} Z_i^{II})$. From parts (i) and (iv) of the proofs of Theorem 2 and Lemma 1 we have $\operatorname{plim}_{N\to\infty} N^{-1} M_{11} = \sigma^{-2}\tilde{\Sigma}_{55,11} = (\sigma^2/(2\lambda))H$ and $\operatorname{plim}_{N\to\infty} M_{22} = \sigma^2 I$. Moreover, because $\lim_{N\to\infty} \mathrm{E}(y_{i,1}\Delta y_{i,t}) = -\sigma^2/2$ and $\lim_{N\to\infty} N^{-1}\mathrm{Var}(y_{i,1}\Delta y_{i,t}) = \sigma^4/(2\lambda)$, we have $\operatorname{plim}_{N\to\infty} N^{-1/2} M_{12} = \operatorname{plim}_{N\to\infty} N^{-1/2} M_{21}' = 0$.

Note that $W_{N,SYS1} = \begin{bmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{bmatrix}^{-1} = \begin{bmatrix} M^{11} & M^{12} \\ M^{21} & M^{22} \end{bmatrix}$, where $M^{11} = N^{-1}(M_{11}/N - M_{12}M_{22}^{-1}M_{21}/N)^{-1}$, $M^{12} = -M^{11}M_{12}M_{22}^{-1}$, $M^{21} = -M_{22}^{-1}M_{21}M^{11}$, and $M^{22} = M_{22}^{-1} + M_{22}^{-1}M_{21}M^{11}M_{12}M_{22}^{-1}$.

The first claim follows now straightforwardly by Slutsky's theorem. The proof of the second claim is very similar.

Part (ii): Note that $\tilde{Z}_i^{S\prime}[\widehat{\Delta v}_i'\ \widehat{v}_i']' = ((Z_i^{I\prime}\widehat{\Delta v}_i)'\ (\tilde{Z}_i^{II\prime}[\widehat{\Delta v}_i'\ \widehat{v}_i']')')'$.

From part (iv) of Lemma 1 we have $N^{-1}\sum_{i=1}^{N}\tilde{Z}_i^{II\prime}[\widehat{\Delta v}_i'\ \widehat{v}_i']'[\widehat{\Delta v}_i'\ \widehat{v}_i']\tilde{Z}_i^{II} \xrightarrow{d} \sigma^4\tilde{W}_{Arbov2}^{-1}$.

Lemma A.3, our model assumptions, and the Cauchy-Schwarz inequality imply that $\lim_{N\to\infty} N^{-1}\mathrm{E}[(\Delta v_{i,s})^2 u_{i,t}^2] < \infty$ and $\lim_{N\to\infty} N^{-2}\mathrm{E}[(\Delta v_{i,s})^4 u_{i,t}^4] < \infty\ \forall s,t$ and hence $N^{-2}\sum_{i=1}^{N}(Z_i^{I\prime}\Delta v_i \Delta v_i' Z_i^I) \xrightarrow{q.m.} \lim_{N\to\infty} N^{-1}\mathrm{E}(Z_i^{I\prime}\Delta v_i \Delta v_i' Z_i^I)$. In part (i) of the proof of Lemma 1 we showed that $\lim_{N\to\infty} N^{-1}\mathrm{E}(Z_i^{I\prime}\Delta v_i \Delta v_i' Z_i^I) = \lim_{N\to\infty}\sigma^2 N^{-1}\mathrm{E}(Z_i^{I\prime} H Z_i^I) = (\frac{\sigma^4}{2\lambda})H$. It then follows along the lines of part (iv) of the proof of Lemma 1 that $\operatorname{plim}_{N\to\infty} N^{-2}\sum_{i=1}^{N}(Z_i^{I\prime}\widehat{\Delta v}_i\widehat{\Delta v_i'}Z_i^I) = (\frac{\sigma^4}{2\lambda})H$.

Similar arguments show that $\operatorname{plim}_{N\to\infty} N^{-3/2}\sum_{i=1}^{N}(K^{II} Z_i^{II\prime} v_i \Delta v_i' Z_i^I) = 0$ and $\operatorname{plim}_{N\to\infty} N^{-2}\sum_{i=1}^{N}(K^{II} Z_i^{II\prime} y_{i,-1}\Delta v_i' Z_i^I) = \operatorname{plim}_{N\to\infty} N^{-2}\sum_{i=1}^{N}(y_{i,1}^2 K^{II} Z_i^{II\prime}\iota\Delta v_i') = (\sigma^2/2\lambda)K^{II}\mathrm{E}[\operatorname{diag}(\varepsilon_i^2,\ldots,\varepsilon_i^{T-1})'\iota\Delta\varepsilon_i'] = (\sigma^4/2\lambda)K^{II}D$. Note that $K^{II}Z_i^{II\prime}\widehat{v}_i = \tilde{Z}_i^{II\prime}[\widehat{\Delta v}_i'\ \widehat{v}_i']$. It follows that $N^{-3/2}\sum_{i=1}^{N}(\tilde{Z}_i^{II\prime}[\widehat{\Delta v}_i'\ \widehat{v}_i']\widehat{\Delta v_i'}Z_i^I) \xrightarrow{d} -(\sigma^4/2\lambda)\mathcal{K}K^{II}D$.

Noting that $\tilde{W}_{SYS2}^{11} = (1/(2\lambda))H$, $\tilde{W}_{SYS2}^{21} = (\tilde{W}_{SYS2}^{12})' = -(1/(2\lambda))\mathcal{K}K^{II}D$, and $\tilde{W}_{SYS2}^{22} = \tilde{W}_{Arbov2}^{-1}$, we conclude that $\operatorname{diag}(N^{-1}I_{T-2}, N^{-1/2}I_m)\sum_{i=1}^{N}(\tilde{Z}_i^{S\prime}[\widehat{\Delta v_i'}\ \widehat{v}_i']' \times [\widehat{\Delta v_i'}\ \widehat{v}_i']\tilde{Z}_i^S)\operatorname{diag}(N^{-1}I_{T-2}, N^{-1/2}I_m) \xrightarrow{d} \sigma^4\tilde{W}_{SYS2}^{-1}$.

Part (iii): From part (i) of Lemma 1 we have $N^{-1}\sum_{i=1}^{N} Z_i^{I\prime}\Delta v_i \xrightarrow{d} \tilde{X}_{51}$, $N^{-1}\times\sum_{i=1}^{N} Z_i^{I\prime}\Delta y_{i,-1} \xrightarrow{d} \tilde{X}_{61}$, $N^{-1/2}\sum_{i=1}^{N} Z_i^{D\prime}\Delta v_i \xrightarrow{d} \tilde{X}_{52}$, and $N^{-1/2}\sum_{i=1}^{N} Z_i^{D\prime}\Delta y_{i,-1} \xrightarrow{d} \tilde{X}_{62}$, with $\tilde{X}_5 = (\tilde{X}_{51}'\ \tilde{X}_{52}')' \sim N(0, \tilde{\Sigma}_{55})$ and $\tilde{X}_6 = (\tilde{X}_{61}'\ \tilde{X}_{62}')' \sim N(\tilde{\mu}_6, \tilde{\Sigma}_{66})$.

From part (ii) of Lemma 1 we have $N^{-1/2}\sum_{i=1}^{N} Z_i^{L\prime} v_i \xrightarrow{d} \tilde{X}_{71} \sim N(0, \tilde{\Sigma}_{77,22})$ and $N^{-1}\sum_{i=1}^{N} Z_i^{L\prime} y_{i,-1} \xrightarrow{d} \tilde{X}_{81} \sim N((\sigma^2\iota/2), (\sigma^4/(2\lambda))I)$.

Recall that $\tilde{Z}_i^{II\prime}[\Delta v_i'\ v_i']' = K^{II} Z_i^{II\prime} v_i$. Let $K^S = \operatorname{diag}(I_{T-2},\ K^{II})$. It follows that $\tilde{Z}_i^{S\prime}[\Delta v_i'\ v_i']' = ((Z_i^{I\prime}\widehat{\Delta v}_i)'\ (\tilde{Z}_i^{II\prime}[\widehat{\Delta v_i'}\ \widehat{v}_i']')')' = K^S Z_i^{S\prime}[\Delta v_i'\ v_i']'$.

Consider now $K^S W_{N,SYS1}^{-1} K^{S\prime} = N^{-1}\sum_{i=1}^{N} K^S Z_i^{S\prime} A Z_i^S K^{S\prime}$ and $K^S W_{N,SYS1b}^{-1} K^{S\prime} = N^{-1}\sum_{i=1}^{N} K^S Z_i^{S\prime} Z_i^S K^{S\prime}$. From part (i) we obtain $\operatorname{plim}_{N\to\infty} N^{-2}\sum_{i=1}^{N}(Z_i^{I\prime} H Z_i^I) =$

https://doi.org/10.1017/S0266466608090531 Published online by Cambridge University Press

$(\sigma^2/(2\lambda))H$, $\operatorname{plim}_{N\to\infty} N^{-1}\sum_{i=1}^N (K^{II} Z_i^{II\prime} Z_i^{II} K^{II\prime}) = \sigma^2 K^{II} K^{II\prime} = \sigma^{-2}\tilde{\Sigma}_{77}$, and $\operatorname{plim}_{N\to\infty} N^{-3/2}\sum_{i=1}^N (Z_i^{I\prime} C Z_i^{II} K^{II\prime}) = 0$. We also have $\operatorname{plim}_{N\to\infty} N^{-2} \times \sum_{i=1}^N (Z_i^{I\prime} Z_i^I) = (\frac{\sigma^2}{2\lambda})I$.

We conclude from the preceding results that $\operatorname{plim}_{N\to\infty}(\widehat{\rho}_{SYSk} - \rho) = 0$ and $\sqrt{N}(\widehat{\rho}_{SYSk} - \rho) \xrightarrow{d} \tilde{X}_8'\tilde{\Sigma}_{77}^{-1}\tilde{X}_7/\tilde{X}_8'\tilde{\Sigma}_{77}^{-1}\tilde{X}_8$, for $k = 1, 1b$.

It immediately follows from part (ii) and the preceding results that $\operatorname{plim}_{N\to\infty}(\widehat{\rho}_{SYS2} - \rho) = 0$ and $\sqrt{N}(\widehat{\rho}_{SYS2} - \rho) \xrightarrow{d} \tilde{X}_{11}'\tilde{W}_{SYS2}\tilde{X}_{10}/\tilde{X}_{11}'\tilde{W}_{SYS2}\tilde{X}_{11}$ with $\tilde{X}_{10} = (\tilde{X}_{51}'\ \tilde{X}_{52}'\ \tilde{X}_{71}')'$ and $\tilde{X}_{11} = (0\ 0\ \tilde{X}_{81}')'$.

Part (iv): From the results mentioned in part (iii) we conclude that $\widehat{\rho}_{SYS} - \rho \xrightarrow{d} (\tilde{X}_{61}' W_{11}\tilde{X}_{51} + \tilde{X}_{81}' W_{31}\tilde{X}_{51})/\tilde{X}_{12}' W\tilde{X}_{12}$ with $\tilde{X}_{12} = (\tilde{X}_{61}'\ 0\ \tilde{X}_{81}')'$. ■

LEMMA A.4. *Let g, d, c^*, and λ be constants such that $0 < g \le 1$, $0 \le d \le g$, $0 < c^* < \infty$, and $0 < \lambda < \infty$. Furthermore, let $\rho = 1 - \lambda N^{-g}$. Then $\lim_{N\to\infty, S/N^d \to c^*} N^{g-d}(1 - \rho^{2(1+S)}) \equiv \hat{q}(\lambda, c^*) < \infty$, where $\hat{q}(\lambda, c^*) = 2\lambda(S+1)$ if $d = 0$; $\hat{q}(\lambda, c^*) = 2\lambda c^*$ if $0 < d < g$; and $\hat{q}(\lambda, c^*) = q(\lambda, c^*)$ if $d = g$.*

Proof of Lemma A.4. Note that $(1 - \rho^{2(1+S)})/(1 - \rho) = (1 + \rho^{1+S})\sum_{k=0}^S \rho^k$ and recall that $\lim_{p\to\infty}(1 + x/p)^p = \exp(x)$ and $q(\lambda, c) = 1 - \exp(-2\lambda c)$. There are three cases: $d = 0$, $0 < d < g$, and $d = g$.

First assume that $d = 0$. Then $S = c^*$ is fixed, and $\lim_{N\to\infty, S/N^d\to c^*}(N^{g-d}(1 - \rho^{2(1+S)})) = \lim_{N\to\infty, S=c^*} \lambda(1 - \rho^{2(1+S)})/(1 - \rho) = \lim_{N\to\infty, S=c^*}(\lambda(1 + \rho^{1+S})\sum_{k=0}^S \rho^k) = 2\lambda(S+1)$.

Next assume that $0 < d < g$. Then we have $\lim_{N\to\infty, S/N^d\to c^*} S^{-1}\sum_{k=0}^S \rho^k = 1$, because $1 \ge \lim_{N\to\infty, S/N^d\to c^*} S^{-1}\sum_{k=0}^S \rho^k \ge \lim_{N\to\infty, S/N^d\to c^*} \rho^S = 1$. It follows that $\lim_{N\to\infty, S/N^d\to c^*} N^{g-d}(1 - \rho^{2(1+S)}) = \lim_{N\to\infty, S/N^d\to c^*}(\lambda N^{-d}(1 - \rho^{2(1+S)})/ \times (1 - \rho)) = \lim_{N\to\infty, S/N^d\to c^*} \lambda(1 + \rho^{1+S})c^* S^{-1}\sum_{k=0}^S \rho^k = 2\lambda c^*$.

Finally assume that $d = g$. Then we have $\lim_{N\to\infty, S/N^d\to c^*} N^{g-d}(1 - \rho^{2(1+S)}) = \lim_{N\to\infty, S/N^g\to c^*}(1 - \rho^{2(1+S)}) = 1 - \exp(-2\lambda c^*) = q(\lambda, c^*) < \infty$. ■

A.4. Proof of Theorem 4. Note that $\lim_{\rho\uparrow 1} \operatorname{Var}(y_{i,-S} - \mu_i) = 0$. Then we obtain for the parameter sequence $\rho = 1 - \lambda N^{-g}$ that $\lim_{N\to\infty, S/N^d\to c^*} N^{-d}\operatorname{Var}(y_{i,1} - \mu_i) = \lim_{N\to\infty, S/N^d\to c^*} \sigma^2 N^{-d}(1 - \rho^{2(1+S)})/(1 - \rho^2) = \lim_{N\to\infty, S/N^d\to c^*} \sigma^2 N^{g-d}(1 - \times \rho^{2(1+S)})/(2\lambda) = \sigma^2\hat{q}(\lambda, c^*)/(2\lambda)$ by Lemma A.4. It follows that $\lim_{N\to\infty, S/N^d\to c^*} N^{-d}\operatorname{Var}(y_{i,1}\Delta y_{i,2}) = \lim_{N\to\infty, S/N^d\to c^*} N^{-d}\operatorname{Var}(y_{i,1}\varepsilon_{i,2}) = \sigma^2(\sigma^2\hat{q}(\lambda, c^*)/(2\lambda) + 1\{d = 0\}\sigma_{-S}^2)$. Furthermore, if $0 < d \le g$, we have $\lim_{N\to\infty, S/N^d\to c^*}[N^{g-d} \times \operatorname{E}(y_{i,1}\Delta y_{i,2})] = -\lim_{N\to\infty, S/N^d\to c^*} \lambda N^{-d}\operatorname{Var}(y_{i,1} - \mu_i) = -\sigma^2\hat{q}(\lambda, c^*)/2$. Finally, if $d = 0$, we obtain that $\lim_{N\to\infty, S/N^d\to c^*} N^{g-d}\operatorname{E}(y_{i,1}\Delta y_{i,2}) = \lambda(\sigma_{\mu y} - \sigma_y^2)$. Note that when $d = 0$, $\sigma_y^2 = \sigma^2\hat{q}(\lambda, c^*)/(2\lambda) + \sigma_{-S}^2$.

When $T = 3$, $\widehat{\rho}_{AB} = \rho + (\sum_{i=1}^N y_{i,1}\Delta\varepsilon_{i,3})/(\sum_{i=1}^N y_{i,1}\Delta y_{i,2})$ and $\widehat{\rho}_{Arbov} = \rho + (\sum_{i=1}^N (\varepsilon_{i,3} + (1 - \rho)\mu_i)\Delta y_{i,2})/(\sum_{i=1}^N y_{i,2}\Delta y_{i,2})$.

https://doi.org/10.1017/S0266466608090531 Published online by Cambridge University Press

Define $X_1 = N^{(-1/2)(1+d)} \sum_{i=1}^N y_{i,1} \Delta\varepsilon_{i,3}$, $X_3 = N^{-1/2} \sum_{i=1}^N ((\varepsilon_{i,3} + (1-\rho)\mu_i)\Delta y_{i,2})$, and $X_4 = N^{-1} \sum_{i=1}^N y_{i,2} \Delta y_{i,2}$. Moreover, let $X_2 = N^{-1/2(1+d)} \sum_{i=1}^N y_{i,1} \Delta y_{i,2}$ if $\Delta_{gd} \leq 0$ and $X_2 = N^{-1+g-d} \sum_{i=1}^N y_{i,1} \Delta y_{i,2}$ if $\Delta_{gd} > 0$.

Recall that $q(\lambda, c) = 0$ if $c = 0$. Then using results similar to those in Lemma A.2 we obtain as $N, S \to \infty$ with $S/N \to c \geq 0$, $S/N^g \to c^+ \geq 0$, and $S/N^d \to c^* > 0$, where $0 < g \leq 1$ and $0 \leq d \leq g$, for the parameter sequence $\rho = 1 - \lambda N^{-g}$ that $X_1 \xrightarrow{d} N(0, 2\sigma^2(\sigma^2 \hat{q}(\lambda, c^*)/(2\lambda) + 1\{d = 0\}\sigma_{-S}^2))$, $X_3 \xrightarrow{d} N(0, \sigma^4)$, and $X_4 \xrightarrow{d} N(\sigma^2(1 - q(\lambda, c^+)/2), \sigma^4 q(\lambda, c)/(2\lambda))$. In addition, we obtain that

$$X_2 \xrightarrow{d} N(\lambda(\sigma_{\mu y} - \sigma_y^2), \sigma^2(\sigma^2 \hat{q}(\lambda, c^*)/(2\lambda) + \sigma_{-S}^2)) \quad \text{if } \Delta_{gd} = 0 \quad \text{and} \quad d = 0,$$

$$X_2 \xrightarrow{d} N(-\sigma^2 \hat{q}(\lambda, c^*)/2, \sigma^4 \hat{q}(\lambda, c^*)/(2\lambda)) \quad \text{if } \Delta_{gd} = 0 \quad \text{and} \quad d > 0,$$

$$X_2 \xrightarrow{d} N(0, \sigma^2(\sigma^2 \hat{q}(\lambda, c^*)/(2\lambda) + 1\{d = 0\}\sigma_{-S}^2)) \quad \text{if } \Delta_{gd} < 0,$$

$$X_2 \xrightarrow{d} \lambda(\sigma_{\mu y} - \sigma_y^2) \quad \text{if } \Delta_{gd} > 0 \quad \text{and} \quad d = 0, \text{ and}$$

$$X_2 \xrightarrow{d} -\sigma^2 \hat{q}(\lambda, c^*)/2 \quad \text{if } \Delta_{gd} > 0 \quad \text{and} \quad d > 0,$$

and we obtain that $\lim_{N\to\infty, S/N^d \to c^*} \mathrm{Cov}(X_1, X_2) = -\sigma^2(\sigma^2 \hat{q}(\lambda, c^*)/(2\lambda) + 1\{d = 0\}\sigma_{-S}^2)$ if $\Delta_{gd} \leq 0$, $\lim_{N\to\infty, S/N^d \to c^*} \mathrm{Cov}(X_1, X_2) = 0$ if $\Delta_{gd} > 0$, and $\lim_{N\to\infty, S/N^d \to c^*} \mathrm{Cov}(X_3, X_4) = 0$.

Parts (i) and (ii) of Theorem 4 follow now straightforwardly from the preceding results. ■

A.5. Proof of Theorem 7. When $T = 3$ and $\rho = 1$, $\widehat{\rho}_{AB} = \rho + (\sum_{i=1}^N y_{i,1} \Delta\varepsilon_{i,3})/(\sum_{i=1}^N y_{i,1}\varepsilon_{i,2})$ and $\widehat{\rho}_{Arbov} = \rho + (\sum_{i=1}^N \varepsilon_{i,3}\varepsilon_{i,2})/(\sum_{i=1}^N y_{i,2}\varepsilon_{i,2})$.

Let us define $X_1 = N^{-1/2(1+d)} \sum_{i=1}^N y_{i,1} \Delta\varepsilon_{i,3}$, $X_2 = N^{-1/2(1+d)} \sum_{i=1}^N y_{i,1}\varepsilon_{i,2}$, $X_3 = N^{-1/2} \sum_{i=1}^N \varepsilon_{i,3}\varepsilon_{i,2}$, and $X_4 = N^{-1/2(1+\bar{d})} \sum_{i=1}^N y_{i,2}\varepsilon_{i,2}$. Let $\sigma_{ur}^4 = \sigma^2(c^*\sigma^2 + 1\{d = 0\}(\sigma_{-S}^2 + \sigma^2))$. Then it is easily seen that as $N, S \to \infty$ with $S/N^d \to c^* > 0$ and $S/N^{\bar{d}} \to c \geq 0$ where $d \geq 0$ and $\bar{d} = \max(d, 1)$, one obtains that $X_1 \xrightarrow{d} N(0, 2\sigma_{ur}^4)$, $X_2 \xrightarrow{d} N(0, \sigma_{ur}^4)$, $X_3 \xrightarrow{d} N(0, \sigma^4)$, and $X_4 \xrightarrow{d} N(1\{d \leq 1\}\sigma^2, c\sigma^4)$. In addition, one obtains that $\lim_{N\to\infty, S/N^d \to c^*} \mathrm{Cov}(X_1, X_2) = -\sigma_{ur}^4$ and $\lim_{N\to\infty, S/N^d \to c^*} \mathrm{Cov}(X_3, X_4) = 0$. Parts (i) and (ii) of Theorem 7 now follow straightforwardly from the preceding results. ■

LEMMA A.5. *Let $\rho = 1$. Furthermore, let $N \to \infty$, $S \to \infty$ simultaneously with $S/N^d \to c^* > 0$, where $d \geq 0$. Let $u_{i,t} = y_{i,t} - \mu_i$, let l, p, and q be positive integers, and let $\kappa_q = \left|\mathrm{E}(\varepsilon_{i,t}^q)/\sigma^q\right| < \infty$. Then assuming that the sum of the powers of $\varepsilon_{i,k}$, $u_{i,s}$, $u_{i,t}$, and $\Delta u_{i,t}$ is less than or equal to 8, we have*

(i) $\lim_{N\to\infty, S/N^d \to \hat{c}} N^{-dp} \mathrm{E}(u_{i,t}^{2p}) < \infty$,
(ii) $\lim_{N\to\infty, S/N^d \to \hat{c}} N^{-dp} \left|\mathrm{E}(u_{i,t}^{2p+1})\right| < \infty$,
(iii) $\lim_{N\to\infty, S/N^d \to \hat{c}} \mathrm{E}[(\Delta u_{i,t})^p] = \kappa_p \sigma^p$,
(iv) $\lim_{N\to\infty, S/N^d \to \hat{c}} N^{-dp} \left|\mathrm{E}[u_{i,s}^{2p}(\Delta u_{i,t})^q]\right| < \infty$,

https://doi.org/10.1017/S0266466608090531 Published online by Cambridge University Press

(v) $\lim_{N\to\infty,S/N^d\to\hat{c}} N^{-dp}\left|\mathrm{E}[u_{i,s}^{2p+1}(\Delta u_{i,t})^q]\right| < \infty$,

(vi) $\lim_{N\to\infty,S/N^d\to\hat{c}} N^{-dp}\left|\mathrm{E}[\varepsilon_{i,k}^l u_{i,s}^{2p}(\Delta u_{i,t})^q]\right| < \infty$, *and*

(vii) $\lim_{N\to\infty,S/N^d\to\hat{c}} N^{-dp}\left|\mathrm{E}[\varepsilon_{i,k}^l u_{i,s}^{2p+1}(\Delta u_{i,t})^q]\right| < \infty$.

Proof of Lemma A.5. Note that $u_{i,t} = u_{i,t-1} + \varepsilon_{i,t} = \sum_{k=-S+1}^{t}\varepsilon_k$ and $\Delta u_{i,t} = \varepsilon_{i,t}$. Note also that $u_{i,-S} = 0$.

When $d = 0$ the proofs for (i)–(vii) are trivial because in this case the moments can be written as sums of a finite number of bounded terms.

Therefore from now on we assume that $d > 0$.

Let $\lim_{N\to\infty}$ be short for $\lim_{N\to\infty,S/N^d\to c^*}$.

We prove (i) and (ii) together.

First consider $p = 0$: $\mathrm{E}(u_{i,t}^0) = 1$ and $\mathrm{E}(u_{i,t}^1) = 0$ and hence $\lim_{N\to\infty}\mathrm{E}(u_{i,t}^0) < \infty$ and $\lim_{N\to\infty}\left|\mathrm{E}(u_{i,t}^1)\right| < \infty$. Now let $p > 0$: $\lim_{N\to\infty} N^{-dp}\mathrm{E}(u_{i,t}^{2p}) = \lim_{N\to\infty}[N^{-dp} \times \sum_{k_1,k_2,\ldots,k_p=-S+1}^{t}\mathrm{E}(\prod_{m=1}^{p}\varepsilon_{i,k_m}^2)] = c^{*p}\sigma^{2p} < \infty$ and $\lim_{N\to\infty} N^{-dp}\left|\mathrm{E}(u_{i,t}^{2p+1})\right| = \lim_{N\to\infty} N^{-dp}\sum_{k_1,k_2,\ldots,k_p=-S+1}^{t}\left(\mathrm{E}\left(\prod_{m=1}^{p-1}\varepsilon_{i,k_m}^2\right)\left|\mathrm{E}\left(\varepsilon_{i,k_p}^3\right)\right|\right) = c^{*p}\,|\kappa_3|\,\sigma^{2p+1} < \infty$.

The proofs of (iii) – (vii) are now straightforward.

(iii) $\lim_{N\to\infty}\mathrm{E}[(\Delta u_{i,t})^p] = \mathrm{E}[(\varepsilon_{i,t})^p] = \kappa_p\sigma^p$,

(iv) $\lim_{N\to\infty} N^{-dp}\left|\mathrm{E}[u_{i,s}^{2p}(\Delta u_{i,t})^q]\right| = \lim_{N\to\infty} N^{-dp}\left|\mathrm{E}[u_{i,s}^{2p}\varepsilon_{i,t}^q]\right| < \infty$,

(v) $\lim_{N\to\infty} N^{-dp}\left|\mathrm{E}[u_{i,s}^{2p+1}(\Delta u_{i,t})^q]\right| = \lim_{N\to\infty} N^{-dp}\mathrm{E}[u_{i,s}^{2p+1}\varepsilon_{i,t}^q] < \infty$,

(vi) $\lim_{N\to\infty} N^{-dp}\left|\mathrm{E}[\varepsilon_{i,k}^l u_{i,s}^{2p}(\Delta u_{i,t})^q]\right| = \lim_{N\to\infty} N^{-dp}\left|\mathrm{E}[\varepsilon_{i,k}^l u_{i,s}^{2p}\varepsilon_{i,t}^q]\right| < \infty$, and finally

(vii) $\lim_{N\to\infty} N^{-dp}\left|\mathrm{E}[\varepsilon_{i,k}^l u_{i,s}^{2p+1}(\Delta u_{i,t})^q]\right| = \lim_{N\to\infty} N^{-dp}\mathrm{E}[\varepsilon_{i,k}^l u_{i,s}^{2p+1}\varepsilon_{i,t}^q] < \infty$.

■

A.6. Proofs of Lemma 3 and Theorem 8. In these proofs (p)$\lim_{N\to\infty}$ is short for (p)$\lim_{N\to\infty,S/N^d\to c^*}$. Note that when $\rho = 1$, $v_{i,t} = \varepsilon_{i,t}$, $\Delta y_{i,t} = \varepsilon_{i,t}$, $\Delta y_i^t = \varepsilon_i^t$, and $Z_i^{II} = \mathrm{diag}(\varepsilon_i^2, \ldots, \varepsilon_i^{T-1})$. ■

Proof of Lemma 3.

Part (i): The proof of the first two results is similar to the proof of the first part of Theorem 7. As $N, S \to \infty$ with $S/N^d \to c^* > 0$ and $S/N^{\bar{d}} \to c \geq 0$ where $d \geq 0$ and $\bar{d} = \max(d, 1)$, we obtain that

$$\begin{bmatrix} N^{(-1/2)(1+d)}\sum_{i=1}^{N} Z_i^{I\prime}\Delta v_i \\ N^{(-1/2)(1+d)}\sum_{i=1}^{N} Z_i^{I\prime}\Delta y_{i,-1} \end{bmatrix} \xrightarrow{d} \begin{bmatrix} \hat{X}_{51} \\ \hat{X}_{61} \end{bmatrix} \sim N\left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \sigma_{ur}^4\begin{pmatrix} H & -C' \\ -C & I \end{pmatrix}\right] \quad \textit{and}$$

$$\begin{bmatrix} N^{-1/2}\sum_{i=1}^{N} Z_i^{D\prime}\Delta v_i \\ N^{-1/2}\sum_{i=1}^{N} Z_i^{D\prime}\Delta y_{i,-1} \end{bmatrix} \xrightarrow{d} \begin{bmatrix} \hat{X}_{52} \\ \hat{X}_{62} \end{bmatrix}$$

$$\sim N\left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \sigma^4\begin{pmatrix} \mathrm{diag}(H_{T-3}, H_{T-4}, \ldots, H_1) & -\mathrm{diag}(C_{T-3}, C_{T-4}, \ldots, C_1)' \\ -\mathrm{diag}(C_{T-3}, C_{T-4}, \ldots, C_1) & I \end{pmatrix}\right].$$

https://doi.org/10.1017/S0266466608090531 Published online by Cambridge University Press

Note that $\hat{\Sigma}_{56,11} = \mathrm{E}(\hat{X}_{51}\hat{X}'_{61}) = -\sigma^4_{ur}C' \neq 0$ and $\hat{\Sigma}_{56,22} = \mathrm{E}(\hat{X}_{52}\hat{X}'_{62}) = -\sigma^4 \times \mathrm{diag}(C_{T-3}, C_{T-4}, \ldots, C_1)' \neq 0$.

Next consider the off-diagonal blocks $\hat{\Sigma}_{55,12}$, $\hat{\Sigma}_{66,12}$, $\hat{\Sigma}_{56,12}$, and $\hat{\Sigma}_{56,21}$. We have $\hat{\Sigma}_{55,12} = \mathrm{E}(\hat{X}_{51}\hat{X}'_{52}) = \lim_{N\to\infty}\sigma^2 N^{(-1/2)d}\mathrm{E}(Z_i^{I\prime}HZ_i^D) = 0$. Similarly, $\hat{\Sigma}_{66,12} = \mathrm{E}(\hat{X}_{61}\hat{X}'_{62}) = \lim_{N\to\infty}\sigma^2 N^{(-1/2)d}\mathrm{E}(Z_i^{I\prime}Z_i^D) = 0$, $\hat{\Sigma}_{56,12} = \mathrm{E}(\hat{X}_{51}\hat{X}'_{62}) = -\lim_{N\to\infty}\sigma^2 N^{(-1/2)d}\mathrm{E}(Z_i^{I\prime}C'Z_i^D) = 0$, and $\hat{\Sigma}_{56,21} = \mathrm{E}(\hat{X}_{52}\hat{X}'_{61}) = -\lim_{N\to\infty}[\sigma^2 N^{(-1/2)d} \times \mathrm{E}(Z_i^{D\prime}C'Z_i^I)] = 0$.

Finally, it is easily seen that if $T = 3$, then $\hat{X}_5 = \hat{X}_{51} = \sigma^2_{ur}\hat{X}_1$ and $\hat{X}_6 = \hat{X}_{61} = \sigma^2_{ur}\hat{X}_2$, where $\hat{X}_1$ and $\hat{X}_2$ are defined in Theorem 7.

Part (ii): The proof of the first two results is similar to the proof of the last part of Theorem 7. As $N, S \to \infty$ with $S/N^d \to c^* > 0$ and $S/N^{\bar{d}} \to c \geq 0$ where $d \geq 0$ and $\bar{d} = \max(d, 1)$, we obtain that

$$\begin{bmatrix} N^{-1/2}\sum_{i=1}^N Z_i^{L\prime}v_i \\ N^{(-1/2)(1+\bar{d})}\sum_{i=1}^N Z_i^{L\prime}y_{i,-1} \end{bmatrix} \xrightarrow{d} \begin{bmatrix} \hat{X}_{71} \\ \hat{X}_{81} \end{bmatrix} \sim N\left[\begin{pmatrix} 0 \\ 1\{d \leq 1\}\sigma^2\iota \end{pmatrix}, \sigma^4\begin{pmatrix} I & 0 \\ 0 & cI \end{pmatrix}\right].$$

Note that $\mathrm{E}(\hat{X}_{71}\hat{X}'_{81}) = 0$. Because $\lim_{N\to\infty} N^{-1/2\bar{d}}\mathrm{E}(Z_i^{D\prime}\Delta v_i y'_{i,-1}Z_i^L) = 0$, we also have $\mathrm{E}(\hat{X}_{52}\hat{X}'_{81}) = 0$. We conclude that $\mathrm{E}(\hat{X}_7\hat{X}'_8) = 0$. Furthermore it is easily verified that $\hat{X}_7 \sim N(0, \hat{\Sigma}_{77})$ with $\hat{\Sigma}_{77} = \lim_{N\to\infty} K^{II}\mathrm{E}(Z_i^{II\prime}v_i v_i'Z_i^{II})K^{II\prime} = \sigma^4 K^{II}K^{II\prime}$.

Finally, it is easily seen that if $T = 3$, then $\hat{X}_7 = \hat{X}_{71} = \sigma^2\hat{X}_3$ and $\hat{X}_8 = \hat{X}_{81} = \sigma^2\hat{X}_4$, where $\hat{X}_3$ and $\hat{X}_4$ are defined in Theorem 7.

Part (iii): Note that $\widehat{v}_i(\widehat{\rho}_1) = y_i - \widehat{\rho}_1 y_{i,-1} = v_i + (\rho - \widehat{\rho}_1)y_{i,-1}$ and $\Delta\widehat{v}_i(\widehat{\rho}_1) = \Delta y_i - \widehat{\rho}_1\Delta y_{i,-1} = \Delta v_i + (\rho - \widehat{\rho}_1)\Delta y_{i,-1}$.

Lemma A.5 implies that $\lim_{N\to\infty}\mathrm{E}(\Delta u_{i,t})^8 < \infty$, $\lim_{N\to\infty} N^{-d}\mathrm{E}[u_{i,s}^2(\Delta u_{i,t})^6] < \infty$, $\lim_{N\to\infty} N^{-2d}\mathrm{E}[u_{i,s}^4(\Delta u_{i,t})^4] < \infty$, $\lim_{N\to\infty}\mathrm{E}[(\Delta u_{i,t})^4] < \infty$, $\lim_{N\to\infty}\left|\mathrm{E}[u_{i,s} \times (\Delta u_{i,t})^3]\right| < \infty$, and $\lim_{N\to\infty} N^{-d}\mathrm{E}[u_{i,s}^2(\Delta u_{i,t})^2] < \infty\ \forall s,t$. It follows from the first result, our model assumptions, and the Cauchy–Schwarz inequality that $\lim_{N\to\infty}\mathrm{Var}([v_{i,s}\Delta y_{i,t}]^2) < \infty\ \forall s,t$ and hence $N^{-1}\sum_{i=1}^N(\tilde{Z}_i^{II\prime}[\Delta v_i'\ v_i']'[\Delta v_i'\ v_i']\tilde{Z}_i^{II}) \xrightarrow{q.m.} \lim_{N\to\infty} K^{II}\mathrm{E}(Z_i^{II\prime}v_i v_i'Z_i^{II})K^{II\prime} = \sigma^4 K^{II}K^{II\prime}$. Similarly it follows that $N^{-(1+\bar{d})}\sum_{i=1}^N(Z_i^{L\prime}y_{i,-1}y'_{i,-1}Z_i^L) \xrightarrow{q.m.} \lim_{N\to\infty} N^{-\bar{d}}\mathrm{E}(Z_i^{L\prime}y_{i,-1}y'_{i,-1}Z_i^L) = c\sigma^4 I$, $N^{-2}\sum_{i=1}^N(Z_i^{D\prime}\Delta y_{i,-1}\Delta y'_{i,-1}Z_i^{D\prime}) \xrightarrow{q.m.} 0$ and $N^{-2}\sum_{i=1}^N(Z_i^{L\prime}y_{i,-1}\Delta y'_{i,-1}Z_i^{D\prime}) \xrightarrow{q.m.} 0$.

Lemma A.5 and our model assumptions also imply that $\lim_{N\to\infty}\left|\mathrm{E}[v_{i,k}u_{i,s} \times (\Delta u_{i,t})^2]\right| < \infty$ and $\lim_{N\to\infty} N^{-d}\mathrm{E}[v_{i,k}^2u_{i,s}^2(\Delta u_{i,t})^4] < \infty\ \forall k,s,t$. It follows from these results, our model assumptions, and the Cauchy–Schwarz inequality that $N^{-3/2}\sum_{i=1}^N(Z_i^{L\prime}v_i y'_{i,-1}Z_i^L) \xrightarrow{q.m.} 0$, $N^{-3/2}\sum_{i=1}^N(Z_i^{L\prime}v_i\Delta y'_{i,-1}Z_i^D) \xrightarrow{q.m.} 0$, $N^{-3/2}\sum_{i=1}^N(Z_i^{L\prime}y_{i,-1} \times \Delta v_i'Z_i^D) \xrightarrow{q.m.} 0$, and $N^{-3/2}\sum_{i=1}^N(Z_i^{D\prime}\Delta y_{i,-1}\Delta v_i'Z_i^D) \xrightarrow{q.m.} 0$.

We conclude from the preceding results that $\mathrm{diag}(N^{-1/2}I_{m-(T-2)}, N^{-1/2\bar{d}}I_{T-2}) \times \sum_{i=1}^N\left(\tilde{Z}_i^{II\prime}[\widehat{\Delta v}_i'(\widehat{\rho}_1)\widehat{v}_i'(\widehat{\rho}_1)]'[\widehat{\Delta v}_i'(\widehat{\rho}_1)\widehat{v}_i'(\widehat{\rho}_1)]\tilde{Z}_i^{II}\right)\mathrm{diag}(N^{-1/2}I_{m-(T-2)}, N^{-1/2\bar{d}}I_{T-2}) \xrightarrow{d} \sigma^4 K^{II}K^{II\prime} + c\sigma^4\mathcal{K}^2\mathrm{diag}(O_{m-(T-2)}, I_{T-2})$; $N^{-1}\sum_{i=1}^N\left(\tilde{Z}_i^{II\prime}[\widehat{\Delta v}_i'(\widehat{\rho}_{1b})\ \widehat{v}_i'(\widehat{\rho}_{1b})]' \times [\widehat{\Delta v}_i'(\widehat{\rho}_{1b})\ \widehat{v}_i'(\widehat{\rho}_{1b})]\tilde{Z}_i^{II}\right) \xrightarrow{d} \sigma^4 K^{II}K^{II\prime} + c\sigma^4\mathcal{K}^2\mathrm{diag}(O_{m-(T-2)}, I_{T-2})$.

https://doi.org/10.1017/S0266466608090531 Published online by Cambridge University Press

Note that $\tilde{Z}_i^{S\prime}[\widehat{\Delta v}_i'\ \widehat{v}_i']' = ((Z_i^{I\prime}\widehat{\Delta v}_i)'\ (\tilde{Z}_i^{II\prime}[\widehat{\Delta v}_i'\ \widehat{v}_i']')')'$.

Lemma A.5, our model assumptions, and the Cauchy–Schwarz inequality imply that $\lim_{N\to\infty} N^{-d}\mathrm{E}[(\Delta v_{i,s})^2 u_{i,t}^2] < \infty$ and $\lim_{N\to\infty} N^{-2d}\mathrm{E}[(\Delta v_{i,s})^4 u_{i,t}^4] < \infty\ \forall s,t$, and hence $N^{-(1+d)}\sum_{i=1}^N (Z_i^{I\prime}\Delta v_i\,\Delta v_i' Z_i^I) \xrightarrow{q.m.} \lim_{N\to\infty} N^{-d}\mathrm{E}(Z_i^{I\prime}\Delta v_i\,\Delta v_i' Z_i^I)$. From part (i) we have that $\lim_{N\to\infty} N^{-d}\mathrm{E}(Z_i^{I\prime}\Delta v_i\,\Delta v_i' Z_i^I) = \lim_{N\to\infty}\sigma^2 N^{-d}\mathrm{E}(Z_i^{I\prime} H Z_i^I) = \sigma_{ur}^4 H$. It follows that $\mathrm{plim}_{N\to\infty} N^{-(1+d)}\sum_{i=1}^N (Z_i^{I\prime}\widehat{\Delta v}_i\,\widehat{\Delta v}_i' Z_i^I) = \sigma_{ur}^4 H$.

Similar arguments show that $\mathrm{plim}_{N\to\infty} N^{(-1/2)(2+d)}\sum_{i=1}^N (K^{II} Z_i^{II\prime} v_i\,\Delta v_i' Z_i^I) = 0$, $\mathrm{plim}_{N\to\infty} N^{(-1/2)(3+d)}\sum_{i=1}^N (Z_i^{D\prime}\Delta y_{i,-1}\,\Delta v_i' Z_i^I) = 0$, and $\mathrm{plim}_{N\to\infty}\Big(N^{(-1/2)(2+d+\bar{d})}\times \sum_{i=1}^N (K^{II} Z_i^{II\prime} y_{i,-1}\,\Delta v_i' Z_i^I)\Big) = \mathrm{plim}_{N\to\infty} N^{(-1/2)(2+d+\bar{d})}\sum_{i=1}^N (y_{i,1}^2 K^{II} Z_i^{II\prime}\iota\Delta v_i') = c\sigma^2 K^{II}\mathrm{E}[\mathrm{diag}(\varepsilon_i^2,\ldots,\varepsilon_i^{T-1})'\iota\Delta\varepsilon_i'] = c\sigma^4 K^{II} D$. Note that $K^{II} Z_i^{II\prime}\widehat{v}_i = \tilde{Z}_i^{II\prime}[\widehat{\Delta v}_i'\ \widehat{v}_i']$. It follows that $N^{(-1/2)(1+d+\bar{d})}\sum_{i=1}^N (\tilde{Z}_i^{II\prime}[\widehat{\Delta v}_i'(\widehat{\rho}_1)\ \widehat{v}_i'(\widehat{\rho}_1)]\widehat{\Delta v}_i'(\widehat{\rho}_1) Z_i^I) \xrightarrow{d} -c\sigma^4\mathcal{K} K^{II} D$ and $N^{(-1/2)(2+d)}\sum_{i=1}^N (\tilde{Z}_i^{II\prime}[\widehat{\Delta v}_i'(\widehat{\rho}_{1b})\ \widehat{v}_i'(\widehat{\rho}_{1b})]\widehat{\Delta v}_i'(\widehat{\rho}_{1b}) Z_i^I) \xrightarrow{d} -c\sigma^4\mathcal{K} K^{II} D$.

We conclude that $\mathrm{diag}(N^{(-1/2)(1+d)} I_{T-2}, N^{-1/2} I_{m-(T-2)}, N^{-1/2\bar{d}} I_{T-2})\sum_{i=1}^N \Big(\tilde{Z}_i^{S\prime} \times [\widehat{\Delta v}_i'(\widehat{\rho}_1)\ \widehat{v}_i'(\widehat{\rho}_1)]'\ [\widehat{\Delta v}_i'(\widehat{\rho}_1)\ \widehat{v}_i'(\widehat{\rho}_1)]\tilde{Z}_i^S\Big)\,\mathrm{diag}(N^{(-1/2)(1+d)} I_{T-2},\ N^{-1/2} I_{m-(T-2)}, N^{-1/2\bar{d}} I_{T-2}) \xrightarrow{d} \sigma^4\hat{W}_{SYS2}^{-1}$ and that $\mathrm{diag}(N^{(-1/2)(1+d)} I_{T-2}, N^{-1/2} I_{m-(T-2)}, N^{-1/2} I_{T-2})\sum_{i=1}^N \tilde{Z}_i^{S\prime}[\widehat{\Delta v}_i'(\widehat{\rho}_{1b})\ \widehat{v}_i'(\widehat{\rho}_{1b})]'[\widehat{\Delta v}_i'(\widehat{\rho}_{1b})\ \widehat{v}_i'(\widehat{\rho}_{1b})]\tilde{Z}_i^S\,\mathrm{diag}(N^{-1/2(1+d)} I_{T-2}, N^{-1/2} I_{m-(T-2)}, N^{-1/2} I_{T-2}) \xrightarrow{d} \sigma^4\hat{W}_{SYS2b}^{-1}$, where $\hat{W}_{SYS2}^{-1}$ and $\hat{W}_{SYS2b}^{-1}$ are PD matrices. ■

Proof of Theorem 8. The proof relies on results in (the proof of) Lemma 3.

Part (i): Recall that $Z_i' = K^{AB} Z_i^{AB\prime}$, where $Z_i^{AB} = [Z_i^I\ \ Z_i^D]$ and $rank(K^{AB}) = m$. Therefore $\widehat{\rho}_{AB1}$ is equal to a GMM estimator that exploits $\mathrm{E}(Z_i^{AB\prime}\Delta v_i) = 0$ and uses the weight matrix $(N^{-1}\sum_{i=1}^N Z_i^{AB\prime} H Z_i^{AB})^{-1}$. Consider the diagonal path limiting behavior of this weight matrix. Since $\lim_{N\to\infty} N^{-2d}\mathrm{Var}(y_{i,1}^2) < \infty$ by Lemma A.5, $\lim_{N\to\infty} N^{-d}\mathrm{Var}(y_{i,1}\Delta y_{i,t}) < \infty$, and $\lim_{N\to\infty}\mathrm{Var}([\Delta y_{i,t}]^2) < \infty$, it is easily seen that $\sigma^2 N^{-(1+d)}\sum_{i=1}^N \Big(Z_i^{I\prime} H Z_i^I\Big) \xrightarrow{q.m.} \hat{\Sigma}_{55,11}$ and $\sigma^2 N^{(-1/2)(2+d)}\sum_{i=1}^N (Z_i^{I\prime} H Z_i^D) \xrightarrow{q.m.} 0$. Moreover, $\sigma^2 N^{-1}\sum_{i=1}^N (Z_i^{D\prime} H Z_i^D) \xrightarrow{q.m.} \hat{\Sigma}_{55,22}$. Therefore $\mathrm{plim}_{N\to\infty}[\sigma^2 N^{-(1+d)}\sum_{i=1}^N (Z_i^{I\prime} H Z_i^I) = \hat{\Sigma}_{55,11}, \mathrm{plim}_{N\to\infty}\sigma^2 N^{-1}\sum_{i=1}^N (Z_i^{D\prime} H Z_i^D) = \hat{\Sigma}_{55,22}$, and $\mathrm{plim}_{N\to\infty}\sigma^2 N^{(-1/2)(2+d)}\sum_{i=1}^N (Z_i^{I\prime} H Z_i^D) = \hat{\Sigma}_{55,12} = 0$. Furthermore $\hat{\Sigma}_{55}$ is PD. It follows from the preceding results and the proof of Lemma 3 that $\widehat{\rho}_{AB1} - \rho \xrightarrow{d} \hat{X}_6'\hat{\Sigma}_{55}^{-1}\hat{X}_5/\hat{X}_6'\hat{\Sigma}_{55}^{-1}\hat{X}_6$. Assuming that $\widehat{\rho}_{AB}$ exploits $\mathrm{E}(Z_i^{AB\prime}\Delta v_i) = 0$, it also follows that $\widehat{\rho}_{AB} - \rho \xrightarrow{d} \hat{X}_{14}'W\hat{X}_{13}/\hat{X}_{14}'W\hat{X}_{14}$ with $\hat{X}_{13} = (\hat{X}_{51}'\ 1\{d=0\}\hat{X}_{52}')'$ and $\hat{X}_{14} = (\hat{X}_{61}'\ 1\{d=0\}\hat{X}_{62}')'$.

Part (ii): It follows in a straightforward manner from the results in the proof of Lemma 3 that $\mathrm{plim}_{N\to\infty}(\widehat{\rho}_{Arbov} - \rho) = 0$ and $N^{(1/2)\bar{d}}(\widehat{\rho}_{Arbov} - \rho) \xrightarrow{d} \hat{X}_8'W\hat{X}_7/\hat{X}_8'W\hat{X}_8$.

Part (iii): From part (i) of Lemma 3 we have $N^{(-1/2)(1+d)}\sum_{i=1}^N Z_i^{I\prime}\Delta v_i \xrightarrow{d} \hat{X}_{51}$, $N^{(-1/2)(1+d)}\sum_{i=1}^N Z_i^{I\prime}\Delta y_{i,-1} \xrightarrow{d} \hat{X}_{61}$, $N^{-1/2}\sum_{i=1}^N Z_i^{D\prime}\Delta v_i \xrightarrow{d} \hat{X}_{52}$, and $N^{-1/2}\sum_{i=1}^N Z_i^{D\prime}\Delta y_{i,-1} \xrightarrow{d} \hat{X}_{62}$, with $\hat{X}_5 = (\hat{X}_{51}'\ \ \hat{X}_{52}')' \sim N(0,\hat{\Sigma}_{55})$ and $\hat{X}_6 = (\hat{X}_{61}'\hat{X}_{62}')' \sim N(\hat{\mu}_6,\hat{\Sigma}_{66})$.

https://doi.org/10.1017/S0266466608090531 Published online by Cambridge University Press

From part (iii) of Lemma 3 we have $N^{-1/2}\sum_{i=1}^{N} Z_i^{L\prime} v_i \xrightarrow{d} \hat{X}_{71} \sim N(0, \hat{\Sigma}_{77,22})$ and $N^{(-1/2)(1+\bar{d})}\sum_{i=1}^{N} Z_i^{L\prime} y_{i,-1} \xrightarrow{d} \hat{X}_{81} \sim N(1\{d \leq 1\}\sigma^2 \iota, c\sigma^4 I)$.

Recall that $\tilde{Z}_i^{S\prime}[\Delta v_i' \; v_i']' = ((Z_i^{I\prime}\widehat{\Delta v}_i)' \; (\tilde{Z}_i^{II\prime}[\widehat{\Delta v}_i' \; \widehat{v}_i']')')' = K^S Z_i^{S\prime}[\Delta v_i' \; v_i']'$, where $\tilde{Z}_i^{II\prime}[\Delta v_i' \; v_i']' = K^{II} Z_i^{II\prime} v_i$ and $K^S = \text{diag}(I_{T-2}, K^{II})$.

Consider now $K^S W_{N,SYS1}^{-1} K^{S\prime} = N^{-1}\sum_{i=1}^{N} K^S Z_i^{S\prime} A Z_i^S K^{S\prime}$ and $K^S W_{N,SYS1b}^{-1} K^{S\prime} = N^{-1}\sum_{i=1}^{N} K^S Z_i^{S\prime} Z_i^S K^{S\prime}$. It is easily verified that $\text{plim}_{N\to\infty} N^{-(1+d)}\sum_{i=1}^{N}(Z_i^{I\prime} H Z_i^I) = (\sigma_{ur}^4/\sigma^2)H$, $\text{plim}_{N\to\infty} N^{-1}\sum_{i=1}^{N}(K^{II} Z_i^{II\prime} Z_i^{II} K^{II\prime}) = \sigma^2 K^{II} K^{II\prime} = \sigma^{-2}\hat{\Sigma}_{77}$, and $\text{plim}_{N\to\infty} N^{(-1/2)(2+d)}\sum_{i=1}^{N}(Z_i^{I\prime} C Z_i^{II} K^{II\prime}) = 0$. We also have $\text{plim}_{N\to\infty}[N^{-(1+d)} \times \sum_{i=1}^{N}(Z_i^{I\prime} Z_i^I)] = (\sigma_{ur}^4/\sigma^2)I$.

We conclude from the preceding results that $\text{plim}_{N\to\infty}(\widehat{\rho}_{SYSk} - \rho) = 0$ and $N^{(1/2)\bar{d}}(\widehat{\rho}_{SYSk} - \rho) \xrightarrow{d} \hat{X}_8' \hat{\Sigma}_{77}^{-1} \hat{X}_7 / \hat{X}_8' \hat{\Sigma}_{77}^{-1} \hat{X}_8$, for $k = 1, 1b$.

It immediately follows from part (iii) of Lemma 3 and the preceding results that $\text{plim}_{N\to\infty}(\widehat{\rho}_{SYS2k} - \rho) = 0$ for $k = 2, 2b$, and $\sqrt{N}(\widehat{\rho}_{SYS2}(\widehat{\rho}_1) - \rho) \xrightarrow{d} \hat{X}_{11}' \hat{W}_{SYS2} \hat{X}_{10} / \hat{X}_{11}' \hat{W}_{SYS2} \hat{X}_{11}$ with $\hat{X}_{10} = (\hat{X}_{51}' \; \hat{X}_{52}' \; 1\{d \leq 1\}\hat{X}_{71}')'$ and $\hat{X}_{11} = (0 \; \hat{X}_8')'$, and $N^{(1/2)\bar{d}}(\widehat{\rho}_{SYS2}(\widehat{\rho}_{1b}) - \rho) \xrightarrow{d} \hat{X}_{11}' \hat{W}_{SYS2b} \hat{X}_{10b} / \hat{X}_{11}' \hat{W}_{SYS2b} \hat{X}_{11}$ with $\hat{X}_{10b} = (\hat{X}_{51}' \; \hat{X}_7')'$.

Part (iv): From the results mentioned in part (iii) we conclude that $N^{(1/2)(\bar{d}-d)}(\widehat{\rho}_{SYS} - \rho) \xrightarrow{d} \hat{X}_{16}' W \hat{X}_{15} / \hat{X}_{16}' W \hat{X}_{16}$ with $\hat{X}_{15} = (\hat{X}_{51}' \; 1\{d = 0\}\hat{X}_7')'$ and $\hat{X}_{16} = (1\{d \geq 1\}\hat{X}_{61}' \; \hat{X}_8')'$, and that if $d < 1$ $\text{plim}_{N\to\infty}(\widehat{\rho}_{SYS} - \rho) = 0$. ■

A.7. Proof of Theorem 9.

Part (i): Recall that $\tilde{Z}_i^{II} = \text{diag}(Z_i^D, Z_i^L)$ and that $\tilde{Z}_i^S = \text{diag}(Z_i^{AB}, Z_i^L)$, where $Z_i^{AB} = [Z_i^I \; Z_i^D]$. Also recall that $K^{II} Z_i^{II\prime} v_i = \tilde{Z}_i^{II\prime}[\Delta v_i' \; v_i']'$ and from the proof of Theorem 3 that $K^S Z_i^{S\prime}[\Delta v_i' \; v_i']' = \tilde{Z}_i^{S\prime}[\Delta v_i' \; v_i']'$ when $K^S = \text{diag}(I_{T-2}, K^{II})$. Similarly, we have $K^{II} Z_i^{II\prime} y_{i,-1} = \tilde{Z}_i^{II\prime}[\Delta y_{i,-1}' \; y_{i,-1}']'$ and $K^S Z_i^{S\prime}[\Delta y_{i,-1}' \; y_{i,-1}']' = \tilde{Z}_i^{S\prime}[\Delta y_{i,-1}' \; y_{i,-1}']'$.

Note that when $\rho = 1$, $v_{i,t} = \varepsilon_{i,t}$, $\Delta y_{i,t} = \varepsilon_{i,t}$, $\Delta y_i^t = \varepsilon_i^t$, and $Z_i^{II} = \text{diag}(\varepsilon_i^2, \ldots, \varepsilon_i^{T-1})$.

Recall the distributional results from parts (i) and (ii) of the proof of Lemma 3 and in particular that $\hat{X}_7 \perp \hat{X}_8$.

Note that $W_{N,Arbov2}(1) = (N^{-1}\sum_{i=1}^{N} Z_i^{II\prime} v_i v_i' Z_i^{II})^{-1}$ and $W_{N,SYS2}(1) = (N^{-1}\sum_{i=1}^{N} Z_i^{S\prime}[\Delta v_i' \; v_i']'[\Delta v_i' \; v_i] Z_i^S)^{-1}$. Consider now $K^{II} W_{N,Arbov2}^{-1}(1) K^{II\prime}$ and $K^S W_{N,SYS2}^{-1}(1) K^{S\prime}$. Recall that $\sigma^4 K^{II} K^{II\prime} = \hat{\Sigma}_{77}$. Then it is easy to verify that

1. $\text{plim}_{N\to\infty, S/N^d \to c^*} N^{-(1+d)}\sum_{i=1}^{N} Z_i^{I\prime} \Delta v_i \Delta v_i' Z_i^I = \sigma_{ur}^4 H$,
2. $\text{plim}_{N\to\infty, S/N^d \to c^*} N^{-1}\sum_{i=1}^{N} K^{II} Z_i^{II\prime} v_i v_i' Z_i^{II} K^{II\prime} = \hat{\Sigma}_{77}$, and
3. $\text{plim}_{N\to\infty, S/N^d \to c^*} N^{(-1/2)(2+d)}\sum_{i=1}^{N} K^{II} Z_i^{II\prime} v_i \Delta v_i' Z_i^I = 0$.

We conclude from the preceding results that $\text{plim}_{N\to\infty, S/N^d \to c^*}(\widehat{\rho}_{Arbov2,R} - 1) = 0$, and $N^{(1/2)\bar{d}}(\widehat{\rho}_{Arbov2,R} - 1) \xrightarrow{d} \hat{X}_8' \hat{\Sigma}_{77}^{-1} \hat{X}_7 / \hat{X}_8' \hat{\Sigma}_{77}^{-1} \hat{X}_8$. Furthermore, we can conclude that $\text{plim}_{N\to\infty,\, S/N^d \to c^*}(\widehat{\rho}_{SYS2,R} - 1) = 0$ and $N^{(1/2)\bar{d}}(\widehat{\rho}_{SYS2,R} - 1) \xrightarrow{d} \hat{X}_8' \hat{\Sigma}_{77}^{-1} \hat{X}_7 / \hat{X}_8' \hat{\Sigma}_{77}^{-1} \hat{X}_8$.

https://doi.org/10.1017/S0266466608090531 Published online by Cambridge University Press

Note that $SE(\widehat{\rho}_{Arbov2,\rho}) = [N^{-1}\sum_{i=1}^{N}(y'_{i,-1}Z_i^{II}K^{II\prime})(K^{II\prime})^{-1}W_{N,Arbov2}(1)(K^{II})^{-1} \times \sum_{i=1}^{N}(K^{II}Z_i^{II\prime}y_{i,-1})]^{-1/2}$ and that $SE(\widehat{\rho}_{SYS2,R}) = [N^{-1}\sum_{i=1}^{N}((\Delta y'_{i,-1}\ y'_{i,-1})Z_i^{S}K^{S\prime})(K^{S\prime})^{-1}W_{N,SYS2}(1)(K^{S})^{-1}\sum_{i=1}^{N}(K^{S}Z_i^{S\prime}(\Delta y'_{i,-1}\ y'_{i,-1})')]^{-1/2}$.

It follows from the preceding results that $[SE(\widehat{\rho}_{Arbov2,R})]^{-2}/N^{\bar{d}} \xrightarrow{d} \hat{X}'_8\hat{\Sigma}_{77}^{-1}\hat{X}_8$ and hence $(\widehat{\rho}_{Arbov2,R}-1)/SE(\widehat{\rho}_{Arbov2,R}) \xrightarrow{d} (\hat{X}'_8\hat{\Sigma}_{77}^{-1}\hat{X}_8)^{-1/2}\hat{X}'_8\hat{\Sigma}_{77}^{-1}\hat{X}_7$. Because $\hat{X}_7 \perp \hat{X}_8$, $(\hat{X}'_8\hat{\Sigma}_{77}^{-1}\hat{X}_8)^{-1/2}\hat{X}'_8\hat{\Sigma}_{77}^{-1}\hat{X}_7|\hat{X}_8 \sim N(0,1)$, and therefore $(\widehat{\rho}_{Arbov2,R}-1)/SE(\widehat{\rho}_{Arbov2,R}) \xrightarrow{d} N(0,1)$. Similarly we obtain that $[SE(\widehat{\rho}_{SYS2,R})]^{-2}/N^{\bar{d}} \xrightarrow{d} \hat{X}'_8\hat{\Sigma}_{77}^{-1}\hat{X}_8$ and $(\widehat{\rho}_{SYS2,R}-1)/SE(\widehat{\rho}_{SYS2,R}) \xrightarrow{d} N(0,1)$.

Part (ii): We first consider the weight matrix and some other fourth-order moments. Note that $\rho = 1-\lambda N^{-1/2}$. Let $u_{i,t} = y_{i,t} - \mu_i\ \forall t$. Then $\lim_{S,N\to\infty,seq}\mathrm{E}[(\Delta u_{i,t})^8] < \infty$, $\lim_{S,N\to\infty,seq}\mathrm{E}[(\Delta u_{i,t})^2] = \sigma^2$, and provided $s \neq t$, $\lim_{S,N\to\infty,seq}\mathrm{E}(\Delta u_{i,s}\Delta u_{i,t}) = 0$. It follows from these results, our model assumptions, and the Cauchy–Schwarz inequality that $\lim_{S,N\to\infty,seq}\mathrm{E}(Z_i^{II\prime}v_iv_i'Z_i^{II}) = \sigma^4 I$ and $\lim_{S,N\to\infty,seq}\mathrm{Var}([v_{i,s}\Delta y_{i,t}]^2) < \infty$ $\forall s,t$. Consequently we obtain that $N^{-1}\sum_{i=1}^{N}Z_i^{II\prime}v_iv_i'Z_i^{II} \xrightarrow{q.m.} \sigma^4 I$ as $(S,N\to\infty)_{seq}$ and hence $\mathrm{plim}_{S,N\to\infty,seq}\ N^{-1}\sum_{i=1}^{N}Z_i^{II\prime}v_iv_i'Z_i^{II} = \sigma^4 I$. Similar arguments show that $\mathrm{plim}_{S,N\to\infty,seq}\ N^{-1}\sum_{i=1}^{N}Z_i^{II\prime}\Delta y_i\Delta y_i'Z_i^{II} = \sigma^4 I$, $\mathrm{plim}_{S,N\to\infty,seq}[N^{-3/2}\sum_{i=1}^{N}(Z_i^{I\prime} \times (\Delta y_i - \Delta y_{i,-1})(\Delta y_i - \Delta y_{i,-1})'Z_i^{I})] = \mathrm{plim}_{S,N\to\infty,seq}\ N^{-3/2}\sum_{i=1}^{N}Z_i^{I\prime}\Delta v_i\Delta v_i'Z_i^{I} = (\frac{\sigma^4}{2\lambda})H$, and $\mathrm{plim}_{S,N\to\infty,seq}\ N^{-5/4}\sum_{i=1}^{N}Z_i^{I\prime}(\Delta y_i - \Delta y_{i,-1})\Delta y_i'Z_i^{II} = \mathrm{plim}_{S,N\to\infty,seq}\ N^{-5/4}\sum_{i=1}^{N}(Z_i^{I\prime}\Delta v_i\ v_i'Z_i^{II}) = 0$.

Next we consider second moments. By Lemma A.2 we have that $\lim_{S,N\to\infty,seq}\mathrm{E}(u_{i,t}\Delta u_{i,t}) = \sigma^2/2\ \forall t$, and Lemma A.3 yields $\lim_{S,N\to\infty,seq}N^{-1}\mathrm{E}(u_{i,t}^2(\Delta u_{i,t})^2) = 0$ $\forall t$. It follows that $N^{-1}\sum_{i=1}^{N}Z_i^{II\prime}y_{i,-1} \xrightarrow{q.m.} (\sigma^2/2)\iota$ as $(S,N\to\infty)_{seq}$ and hence $\mathrm{plim}_{S,N\to\infty,seq}\ N^{-1}\sum_{i=1}^{N}Z_i^{II\prime}y_{i,-1} \equiv \vec{X}_8 = (\sigma^2/2)\iota$. Similarly, we obtain that $\mathrm{plim}_{S,N\to\infty,seq}\ N^{-1}\sum_{i=1}^{N}Z_i^{I\prime}\Delta y_{i,-1} = -(\sigma^2/2)\iota$.

The preceding results imply that both $\mathrm{plim}_{S,N\to\infty,seq}[SE(\widehat{\rho}_{Arbov2,\rho})]^{-2}/N = \vec{X}'_8\vec{X}_8/\sigma^4 = (1/2)^2 m$ and $\mathrm{plim}_{S,N\to\infty,seq}[SE(\widehat{\rho}_{SYS2,R})]^{-2}/N = \vec{X}'_8\vec{X}_8/\sigma^4$.

Because $N^{-3/4}\sum_{i=1}^{N}Z_i^{I\prime}\Delta v_i \xrightarrow{d} \vec{X}_{51} \sim N(0,(\frac{\sigma^4}{2\lambda})H)$ and $N^{-1/2}\sum_{i=1}^{N}Z_i^{II\prime}v_i \xrightarrow{d} \vec{X}_7 \sim N(0,\sigma^4 I)$ as $(S,N\to\infty)_{seq}$, we also obtain that $(\widehat{\rho}_{Arbov2,R}-\rho)/SE(\widehat{\rho}_{Arbov2,R}) \xrightarrow{d} \sigma^{-2}(\vec{X}'_8\vec{X}_8)^{-1/2}\vec{X}'_8\vec{X}_7$ and $(\widehat{\rho}_{SYS2,R}-\rho)/SE(\widehat{\rho}_{SYS2,R}) \xrightarrow{d} \sigma^{-2}(\vec{X}'_8\vec{X}_8)^{-1/2}\vec{X}'_8\vec{X}_7$ as $(S,N\to\infty)_{seq}$.

Note that $\vec{X}_7 \perp \vec{X}_8$. Therefore $\sigma^{-2}(\vec{X}'_8\vec{X}_8)^{-1/2}\vec{X}'_8\vec{X}_7|\vec{X}_8 \sim N(0,1)$. We conclude that $(\widehat{\rho}_{Arbov2,R}-1)/SE(\widehat{\rho}_{Arbov2,R}) \xrightarrow{d} N(-(\lambda/2)\sqrt{m},1)$ and $(\widehat{\rho}_{SYS2,R}-1)/SE(\widehat{\rho}_{SYS2,R}) \xrightarrow{d} N(-(\lambda/2)\sqrt{m},1)$.

Part (iii): We first consider the weight matrix and some other fourth-order moments. Note that $\rho = 1-\lambda N^{-1/2}$ and $d \geq 0$. Let $u_{i,t} = y_{i,t} - \mu_i\ \forall t$. Then $\lim_{N\to\infty,S/N^d\to c^*}\mathrm{E}[(\Delta u_{i,t})^8] < \infty$, $\lim_{N\to\infty,S/N^d\to c^*}\mathrm{E}[(\Delta u_{i,t})^2] = \sigma^2$, and provided $s \neq t$, $\lim_{N\to\infty,S/N^d\to c^*}\mathrm{E}(\Delta u_{i,s}\Delta u_{i,t}) = 0$. It follows from these results, our model assumptions, and the Cauchy–Schwarz inequality that $\lim_{N\to\infty,S/N^d\to c^*}\mathrm{E}(Z_i^{II\prime}v_iv_i'Z_i^{II}) = \sigma^4 I$ and $\lim_{N\to\infty,S/N^d\to c^*}\mathrm{Var}([v_{i,s}\Delta y_{i,t}]^2) < \infty\ \forall s,t$. Consequently we obtain that $N^{-1}\sum_{i=1}^{N}(Z_i^{II\prime}v_i \times v_i'Z_i^{II}) \xrightarrow{q.m.} \sigma^4 I$ and hence $\mathrm{plim}_{N\to\infty,S/N^d\to c^*}$

https://doi.org/10.1017/S0266466608090531 Published online by Cambridge University Press

$N^{-1}\sum_{i=1}^{N} Z_i^{II\prime} v_i v_i' Z_i^{II} = \sigma^4 I$. Similar arguments show that $\operatorname{plim}_{N\to\infty,S/N^d\to c^*} N^{-1}\sum_{i=1}^{N} Z_i^{II\prime}\Delta y_i \Delta y_i' Z_i^{II} = \sigma^4 I$.

Because $\lim_{S,N\to\infty,seq} N^{-1}\mathrm{E}(\varepsilon_{i,k}^{2l} u_{i,s}^4 (\Delta u_{i,t})^{2q}) < \infty$ $\forall k,s,t$ and $\forall l,q \geq 0$ by Lemma A.3, we have a fortiori $\lim_{N\to\infty,S/N^d\to c^*} N^{-1}\mathrm{Var}([y_{i,1}\Delta y_{i,t}]^2) < \infty$ and $\lim_{N\to\infty,S/N^d\to c^*} N^{-1}\mathrm{Var}([y_{i,1}\Delta v_{i,t}]^2) < \infty$ $\forall t$ when $d \geq 0$. Moreover, when $d = 0$, we have $\lim_{N\to\infty,S/N^d\to c^*} \mathrm{Var}([y_{i,1}\Delta y_{i,t}]^2) < \infty$ and $\lim_{N\to\infty,S/N^d\to c^*} \mathrm{Var}([y_{i,1}\Delta v_{i,t}]^2) < \infty$ $\forall t$. It follows that $\operatorname{plim}_{N\to\infty,S/N^d\to c^*} N^{-(1+\underline{d})}\sum_{i=1}^{N} Z_i^{I\prime} \times (\Delta y_i - \Delta y_{i,-1})(\Delta y_i - \Delta y_{i,-1})' Z_i^I = \operatorname{plim}_{N\to\infty,S/N^d\to c^*} N^{-(1+\underline{d})}\sum_{i=1}^{N} Z_i^{I\prime}\Delta v_i \Delta v_i' Z_i^I = (\sigma^2\vec{q}(\lambda,c^*)/(2\lambda) + 1\{d = 0\}\sigma_{-S}^2)\sigma^2 H$ and $\operatorname{plim}_{N\to\infty,S/N^d\to c^*}[N^{-(1+(1/2)\underline{d})} \times \sum_{i=1}^{N} Z_i^{I\prime}](\Delta y_i - \Delta y_{i,-1})\Delta y_i' Z_i^{II} = \operatorname{plim}_{N\to\infty,S/N^d\to c^*} N^{-(1+(1/2)\underline{d})}\sum_{i=1}^{N}(Z_i^{I\prime} \times (\Delta v_i) v_i' Z_i^{II}) = 0$, where $\underline{d} = \min(d, 1/2)$ and $\vec{q}(\lambda,c^*)/(2\lambda) \equiv \lim_{N\to\infty,S/N^d\to c^*} N^{-\underline{d}}(1-\rho^{2(1+S)})/(1-\rho^2)$ (cf. $\hat{q}(\lambda,c^*)/(2\lambda)$ in the proof of Theorem 4). Note that $\vec{q}(\lambda,c^*)/(2\lambda) = S+1$ if $d = 0$; $\vec{q}(\lambda,c^*)/(2\lambda) = c^*$ if $0 < d < 1/2$; $\vec{q}(\lambda,c^*)/(2\lambda) = q(\lambda,c^*)/(2\lambda)$ if $d = 1/2$; and $\vec{q}(\lambda,c^*)/(2\lambda) = 1/(2\lambda)$ if $d > 1/2$.

Next we consider second moments. When $0 < d < 1/2$, we have $\lim_{N\to\infty,S/N^d\to c^*} \mathrm{E}((1-\rho)u_{i,1}^2) = \lim_{N\to\infty,S/N^d\to c^*} \sigma^2(1-\rho^{2(1+S)})/2 = 0$ by part (ii) of Lemma A.4. When $d = 0$, we obtain $\lim_{N\to\infty,S/N^d\to c^*} \mathrm{E}((1-\rho)u_{i,1}^2) = 0$ also. Finally, when $d \geq 1/2$, we have $\lim_{N\to\infty,S/N^d\to c^*} \mathrm{E}((1-\rho)u_{i,1}^2) = \lim_{N\to\infty,S/N^{1/2}\to c,S/N^d\to c^*}[\sigma^2(1-\rho^{2(1+S)})/2] = \sigma^2 q(\lambda,c)/2$.

Recall that if $c = 0$, then $q(\lambda,c) = 0$. Then it is easily verified that for $d \geq 0$ $\lim_{N\to\infty,S/N^d\to c^*} \mathrm{E}(u_{i,t}\Delta u_{i,t}) = \sigma^2 + \lim_{N\to\infty,S/N^d\to c^*} \mathrm{E}(u_{i,t-1}\Delta u_{i,t}) = \sigma^2 + \lim_{N\to\infty,S/N^{1/2}\to c,S/N^d\to c^*} \mathrm{E}((\rho-1)u_{i,1}^2) = \sigma^2 - \sigma^2 q(\lambda,c)/2$. Moreover, because $\lim_{S,N\to\infty,seq} N^{-1}\mathrm{E}(u_{i,t}^2(\Delta u_{i,t})^2) = 0$ by Lemma A.3, we have a fortiori $\lim_{N\to\infty,S/N^d\to c^*} N^{-1}\mathrm{E}(u_{i,t}^2(\Delta u_{i,t})^2) = 0$ when $d \geq 0$. It follows that $N^{-1}\sum_{i=1}^{N} Z_i^{II\prime} y_{i,-1} \xrightarrow{q.m.} \sigma^2(1 - q(\lambda,c)/2)\iota$ and hence $\operatorname{plim}_{N\to\infty,S/N^d\to c^*} N^{-1}\sum_{i=1}^{N} Z_i^{II\prime} y_{i,-1} \equiv \vec{X}_8 = \sigma^2(1 - q(\lambda,c)/2)\iota$. Similarly, we obtain that $\operatorname{plim}_{N\to\infty,S/N^d\to c^*} N^{-(1+(1/2)\underline{d})}\sum_{i=1}^{N} Z_i^{I\prime}\Delta y_{i,-1} = 0$ when $d \geq 0$.

Finally, note that $N^{-1/2(1+\underline{d})}\sum_{i=1}^{N} Z_i^{I\prime}\Delta v_i \xrightarrow{d} \vec{X}_{51} \sim N(0, (\sigma^2\vec{q}(\lambda,c^*)/(2\lambda) + 1\{d = 0\}\sigma_{-S}^2)\sigma^2 H)$ and $N^{-1/2}\sum_{i=1}^{N} Z_i^{II\prime} v_i \xrightarrow{d} \vec{X}_7 \sim N(0,\sigma^4 I)$ as $N,S\to\infty$ with $S/N^{1/2} \to c \geq 0$ and $S/N^d \to c^* > 0$, where $d \geq 0$.

The preceding results imply that $\operatorname{plim}_{N\to\infty,S/N^d\to c^*}[SE(\widehat{\rho}_{Arbov2,R})]^{-2}/N = \vec{X}_8'\vec{X}_8/\sigma^4 = (1-q(\lambda,c)/2)^2 m$ and $\operatorname{plim}_{N\to\infty,S/N^d\to c^*}[SE(\widehat{\rho}_{SYS2,R})]^{-2}/N = \vec{X}_8'\vec{X}_8/\sigma^4$. Moreover, as $N,S\to\infty$ with $S/N^{1/2}\to c \geq 0$ and $S/N^d \to c^* > 0$, where $d \geq 0$, $(\widehat{\rho}_{Arbov2,R} - \rho)/SE(\widehat{\rho}_{Arbov2,R}) \xrightarrow{d} \sigma^{-2}(\vec{X}_8'\vec{X}_8)^{-1/2}\vec{X}_8'\vec{X}_7$ and $(\widehat{\rho}_{SYS2,R} - \rho)/SE(\widehat{\rho}_{SYS2,R}) \xrightarrow{d} \sigma^{-2}(\vec{X}_8'\vec{X}_8)^{-1/2}\vec{X}_8'\vec{X}_7$.

Note that $\vec{X}_7 \perp \vec{X}_8$ Therefore $\sigma^{-2}(\vec{X}_8'\vec{X}_8)^{-1/2}\vec{X}_8'\vec{X}_7|\vec{X}_8 \sim N(0,1)$. We conclude that $(\widehat{\rho}_{Arbov2,R} - 1)/SE(\widehat{\rho}_{Arbov2,R}) \xrightarrow{d} N(-\lambda(1-q(\lambda,c)/2)\sqrt{m},1)$ and $(\widehat{\rho}_{SYS2,R} - 1)/SE(\widehat{\rho}_{SYS2,R}) \xrightarrow{d} N(-\lambda(1-q(\lambda,c)/2)\sqrt{m},1)$. ■

https://doi.org/10.1017/S0266466608090531 Published online by Cambridge University Press