Bias Reduction by Recursive Mean Adjustment in Dynamic Panel Data Models

Chi-Young Choi
University of New Hampshire

Nelson C. Mark
University of Notre Dame and NBER

Donggyu Sul
University of Auckland

July 2, 2004

Abstract

Accurate estimation of the dominant root of a stationary but persistent time series are required to determine the speed at which economic time series, such as real exchange rates or interest rates, adjust towards their mean values. In practice, accuracy is hampered by downward small-sample bias. Recursive mean adjustment has been found to be a useful bias reduction strategy in the regression context. In this paper, we study recursive mean adjustment in dynamic panel data models. When there exists cross-sectional heterogeneity in the dominant root, the recursive mean adjusted SUR estimator is appropriate. When homogeneity restrictions can be imposed, a pooled recursive mean adjusted GLS estimator with fixed effects is the desired estimator. Application of these techniques to a small panel of five eurocurrency rates finds that these interest rates are unit root nonstationary as the bias-corrected autoregressive coefficient exceeds 1.

Keywords: Small sample bias, Recursive mean adjustment, Panel Data, Cross-Sectional Dependence, Interest rate dyanamics

JEL Classification: C23,G10

Introduction

Accurate estimation of the autocorrelation coefficient of a stationary but persistent first-order autoregressive time series $y_t = c + \rho y_{t-1} + u_t$, where $u_t \overset{iid}{\sim} (0, \sigma_u^2)$, must take account of the downward small-sample bias that induced by running the regression with a constant. To see the source of this bias, think of running least squares without a constant on observations that are deviations from the sample mean. Then for any observation t, the regression error u_t is correlated with current and future values of $y_t, y_{t+1}, \ldots y_T$. These values are embedded in the sample mean which is now a component of the explanatory variable.[1] The use of panel data does not eliminate this small sample bias. Nickell (1981), showed that the least squares dummy variable (LSDV) estimator for the dynamic panel regression model with cross-sectionally independent observations is more efficient than OLS but retains some downward bias even as the number of cross-sectional units goes to infinity.

Recursive mean adjustment is a bias reduction technique that has been found to be useful in the regression context [So and Shin (1999)]. The idea is to adjust the observations not with the sample mean but with the recursive mean, $(t-1)^{-1}\sum_{j=1}^{t-1} y_j$. Because it does not contain future values of y_t, one obtains an unbiased control for the constant which orthogonal to the regression error.

In this paper, we study and apply recursive mean adjustment to dynamic models for panel data. The first estimator that we propose is a recursive mean adjusted seemingly unrelated regression estimator, which is appropriate if the autoregressive coefficients across the different time series in the panel are heterogeneous. In order to pool (impose homogeneity across the cross-sectional units), the panel must pass a test of homogeneity restrictions. However, such a test will suffer from size distortion if it is based on a biased estimator. The recursive mean adjusted seemingly unrelated regression estimator has familiar asymptotic properties and achieves favorable reduction of small sample bias and that the associated test of homogeneity restrictions is reasonably sized.

Assuming that homogeneity restrictions can be imposed, sharper results can be obtained by estimating under these restrictions. Here, we propose a pooled recursive mean adjusted generalized least squares estimator with fixed effects. The asymptotic distribution theory for this estimator also works reasonably well in small samples and a notable contribution of the paper is that this recursive mean adjusted generalized least squares estimator can be directly applied estimate more general AR(p) models. This is notable because alternative mean (median) unbiased estimators are generally unavailable in the panel AR(p) case.

The methods we discuss have a wide range of applicability. The use of panel data

[1] Mariott and Pope (1954) and Kendall (1954) discuss and characterize the first-order approximaiton of this bias. Several bias correction strategies have been suggested in the literature. These include mean unbiased estimation, median unbiased estimation, and recursive mean adjustment strategies. See Andrews (1993) for the exactly median unbiased estimator, Andrews and Chen (1994) for approximately median unbiased estimation, and Phillips and Sul (2003a) for mean unbiased estimation.

has become increasingly important in areas such as finance and international macroeconomics where the research focus is on estimating the speed of adjustment of price or financial variables. In international economics, panel data has been used to estimate the speed of adjustment towards purchasing power parity [Frankel and Rose (1996), Papell (2004), Murray and Papell (2002)] whereas in finance research has emphasized the adjustment speed of interest rate dynamics [Chan et. al. (1992), Ball and Torus (1996), Wu and Zhang (1996).]. Previous estimates that did not control for the downward bias [Gibbons and Ramaswamy (1993), Perarson and Sun (1994)] found adjustment speeds to be too fast to be consistent with the observed dynamics in bond yield data. When we apply the methods to estimate the dynamics of a small panel of Eurocurrency interest rates, we find that our bias adjusted estimates imply a zero speed of adjustment. That is, short-term nominal interest rates are not mean reverting, but are unit-root nonstationary.

The remainder of the paper is organized as follows. The next section develops the recursive seemingly unrelated regression estimator which is appropriate under cross-sectional heterogeneity and which provides a convenient framework for testing homogeneity restrictions. Section 2 develops the recursive mean adjusted generalized least squares estimator for panel data with fixed effects. Section 3 reports the results of Monte Carlo experiments that compares performance across alternative estimators. In Section 4, we apply the recursive mean adjusted generalized least squares to estimate the dynamics of interest rates using a small panel of Eurocurrency rates, and Section 5 concludes.

Because we discuss several different estimators with cumbersome names, we use the following labels: RLS–recursive mean adjusted least squares estimator; RSUR–recursive mean adjusted seemingly unrelated regression estimator; RLSDV–recursive mean adjusted least squares dummy variable estimator; RGLS–recursive mean adjusted generalized least squares estimator with fixed effects; MLSDV–mean unbiased least squares dummy variable estimator; MGLS–mean unbiased generalized least squares estimator with fixed effects. In addition, feasible multivariate estimators that employ parametric (nonparametric) estimates of error covariance matrices are indicated accordingly. The appendix contain the details of several arguments made in the text.

1 Recursive mean adjustment under heterogeneity

To fix ideas, we review recursive mean adjustment in the univariate regression context. Subsection 1.2 extends the strategy to the dynamic panel regression under parameter heterogeneity and introduces the RSUR estimator. The strategy is to use RLS to obtain bias corrected residuals in estimation of the error-covariance matrix for RSUR. The RSUR framework also provides a convenient test of homogeneity restrictions across cross-sectional units.

1.1 Recursive Mean Adjusted Least Squares (RLS)

Assume that the time series $\{y_t\}$ is generated by the AR(1) process

$$y_t = c + \rho y_{t-1} + u_t, \tag{1}$$

where $|\rho| \leq 1$, $u_t \overset{iid}{\sim} (0, \sigma_u^2)$, $c = (1-\rho)\mu$, and $E(y_t) = \mu$. Estimating (1) by OLS is equivalent to running OLS without a constant on deviations from the sample mean,

$$\left(y_t - \frac{1}{T}\sum_{j=1}^{T} y_j\right) = \rho\left(y_{t-1} - \frac{1}{T}\sum_{j=1}^{T} y_j\right) + \left(u_t - (1-\rho)\frac{1}{T}\sum_{j=1}^{T} y_j\right). \tag{2}$$

When $0 < \rho < 1$, u_t is positively correlated with current and future values $(y_t, y_{t+1}, \ldots y_T)$, which themselves are contained in the sample mean and it is this correlation that causes the downward bias in OLS.[2]

Instead of adjusting the observations with the full sample mean, the recursive mean adjustment method adjusts each observation y_t with the recursive sample mean, $\mu_{t-1} = (t-1)^{-1}\sum_{j=1}^{t-1} y_j$,

$$\left(y_t - \mu_{t-1}\right) = c + \rho\left(y_{t-1} - \mu_{t-1}\right) + \left(u_t - (1-\rho)\mu_{t-1}\right).$$

Recursive mean adjustment controls for the presence of the constant since $E(1-\rho)\mu_{t-1} = c$ and the explanatory variable is orthogonal to u_t as the recursive mean now contains observations only up through date $t-1$. Recursive mean adjustment in the regression context has been studied by So and Shin (1999). We now extend this strategy to an environment of systems of dynamic regression equations.

1.2 Recursive Mean Adjusted Seemingly Unrelated Regression (RSUR)

Consider a system of dynamic regression equations indexed by $i = 1, \ldots N$. Each time series in the system $\{y_{it}\}$ follows an AR(1) with autocorrelation coefficient ρ_i and unknown constant c_i,

$$y_{it} = c_i + \rho_i y_{it-1} + u_{it}, \tag{3}$$

[2]The correlation between the regressor $\left(y_{t-1} - \frac{1}{T}\sum_{j=1}^{T} y_j\right)$ and the second part of the composite error term $-(1-\rho)\frac{1}{T}\sum_{j=1}^{T} y_j$ is inconsequential for values of ρ near the borderline of stationarity. To see this, note that $E(1-\rho)\left(\frac{1}{T}\sum_{j=1}^{T} y_j\right)\left(y_{t-1} - \mu - \frac{1}{T}\sum_{j=1}^{T}(y_j - \mu)\right) = \frac{(1-\rho)}{T}\left(1 + \sum_{j=1}^{t-1}\rho^j + \sum_{j=1}^{T-t+1}\rho^j\right) - (1-\rho)Var\left(\frac{1}{T}\sum_{j=1}^{T} y_j\right)$ $= \frac{1+\rho-\rho^t-\rho^{T-t+1}}{T} - (1-\rho)\frac{1}{T}\left(\sum_{j=0}^{T-1}\frac{T-j}{T}\rho^j\right)$. For $t = 3$, $T = 30$, $\rho = 0.98$, this expectation is 0.008.

where $|\rho_i| \leq 1$ and the regression errors, which may be cross sectionally correlated, are generated by the single factor structure[3]

$$u_{it} = \delta_i \theta_t + \epsilon_{it}. \tag{4}$$

The factor $\theta_t \overset{iid}{\sim} (0, 1)$ is a time t common shock that may have a differential impact (δ_i) across individuals. The idiosyncratic error $\{\epsilon_{it}\}$ has mean zero and is not cross-sectionally or temporally correlated but may be heteroskedastic across individuals with covariance matrix $\mathbf{\Sigma} = E(\epsilon_{1t}, \ldots \epsilon_{Nt})'(\epsilon_{1t}, \ldots \epsilon_{Nt}) =$diag$[\sigma_1^2, \ldots \sigma_N^2]$.

The RSUR estimator for the system of AR(1) equations with unknown constant is constructed as follows.

Step 1 Estimate each equation $i = 1, \ldots N$ by RLS. Construct the recursive means of the observations, $\mu_{it} = \frac{1}{t-1} \sum_{s=1}^{t-1} y_{is}$, and estimate $y_{it} - \mu_{it} = \rho_i (y_{it-1} - \mu_{it}) + e_{it}$ by OLS. Call the RLS estimator $\rho_{\mathrm{RLS},i}$.

Step 2 Construct the RLS residuals $\hat{u}_{it} = y_{it} - \rho_{\mathrm{RLS},i} y_{it-1} - \widehat{c}_i$, (not $\widehat{e}_{it}$) where $\widehat{c}_i = \frac{1}{T} \sum_{t=1}^{T} \left(y_{it} - \rho_{\mathrm{RLS},i} y_{it-1} \right)$.[4] Let $\underline{\hat{u}}_t = (\hat{u}'_{1t}, \cdots, \hat{u}'_{Nt})'$ and form the sample error covariance matrix,

$$\mathbf{M}_T = \frac{1}{T} \sum_{t=1}^{T} \underline{\hat{u}}_t \underline{\hat{u}}'_t. \tag{5}$$

Let $\underline{\delta} = (\delta_1, \cdots \delta_N)'$. Estimate δ_i and $\mathbf{\Sigma}$ by the iterative method of moments as suggested by Phillips and Sul (2003a).

$$\left(\underline{\hat{\delta}}, \hat{\mathbf{\Sigma}} \right) = \arg \min_{\underline{\delta}, \mathbf{\Sigma}} \left\{ tr \left[\left(\mathbf{M}_T - \mathbf{\Sigma} - \underline{\delta \delta'} \right) \left(\mathbf{M}_T - \mathbf{\Sigma} - \underline{\delta \delta'} \right)' \right] \right\}. \tag{6}$$

Use the estimates to form a parametric estimate of the error covariance matrix,

$$\hat{\mathbf{V}}_u = \hat{\mathbf{\Sigma}} + \underline{\hat{\delta} \hat{\delta}}'. \tag{7}$$

Step 3 Let $z_{it-1} = (y_{it-1} - \mu_{it})$, $\mathbf{Z}_{t-1} =$diag$[z_{1t}, z_{2t}, \ldots z_{Nt}]$, $\underline{Y}_t = (y_{1t} - \mu_{1t}, \ldots, y_{Nt} - \mu_{Nt})'$,

---

[3] The factor structure has been employed by Bai and Ng (2002), Moon and Perron (2003), and Phillips and Sul (2003a). The single factor assumption is not restrictive as our results go through with a multi-factor model.

[4] The RLS residual $\hat{e}_{it}$ contains the recursive mean adjustment term μ_{it}. As a result, the asymptotic variance $\hat{e}_{it}$ is not same as that of u_{it}. That is why one should use $\hat{u}_{it}$ rather than $\hat{e}_{it}$. Note that the composite error term is $e_{it} = [a_i - (1 - \rho_i) \mu_{it}] + u_{it} = -(1 - \rho_i) \frac{1}{t-1} \sum_{s=1}^{t-1} x_{is} + u_{it}$. For any finite T, $Ee'_{it} e_{it} \neq Eu'_{it} u_{it}$ but the difference between the two becomes vanishingly small with increasing sample size. We show in the appendix that $E(e'_{it} e_{it} - u'_{it} u_{it}) = O\left(T^{-1} \ln T\right)$.

and $\underline{\beta} = (\rho_1, \ldots \rho_N)'$. Then the RSUR estimator is

$$\underline{\beta}_{\mathrm{RSUR}} = \left(\sum_{t=1}^{T} \mathbf{Z}_{t-1}' \hat{\mathbf{V}}_u^{-1} \mathbf{Z}_{t-1} \right)^{-1} \left(\sum_{t=1}^{T} \mathbf{Z}_{t-1}' \hat{\mathbf{V}}_u^{-1} \underline{y}_t \right). \tag{8}$$

Note that step 2 produces a *parametric* estimate of the covariance matrix. An alternative approach is to estimate the error covariance matrix $\mathbf{V}_u$ by *nonparametric* methods, say by using $\mathbf{M}_T$. If the parametric structure is known and can be properly modeled by a factor structure, the parametric method is likely to be more efficient in small samples.[5]

For fixed N, as $T \to \infty$ RSUR is asymptotically normally distributed, $\sqrt{T}\left[\underline{\beta}_{\mathrm{RSUR}} - \underline{\beta}\right] \xrightarrow{D} N(0, \mathbf{\Omega})$, where $\mathbf{\Omega} = \lim_{T\to\infty} \left(\frac{1}{T} \sum_{t=1}^{T} \mathbf{Z}_{t-1}' \mathbf{V}_u^{-1} \mathbf{Z}_{t-1} \right)^{-1}$, and the Wald test provides a convenient test of homogeneity restrictions across equations. To test the null hypothesis $H_o : \rho_1 = \rho_2 = \cdots = \rho_N$, form the restriction matrix

$$\mathbf{R} = \begin{bmatrix} 1 & -1 & 0 & \cdots & \cdots & 0 \\ 0 & 1 & -1 & \cdots & \cdots & 0 \\ \vdots & & & & & \vdots \\ 0 & 0 & 0 & \cdots & 1 & -1 \end{bmatrix}. \tag{9}$$

Then under the null hypothesis,

$$W = \left[\mathbf{R}\underline{\beta}_{\mathrm{RSUR}}\right]' \left(\mathbf{R}\widehat{\mathbf{\Omega}}\mathbf{R}'\right)^{-1} \left[\mathbf{R}\underline{\beta}_{\mathrm{RSUR}}\right] \sim \chi^2_{N-1}$$

where $\widehat{\mathbf{\Omega}} = \left(\sum_{t=1}^{T} \mathbf{Z}_t' \hat{\mathbf{V}}_u^{-1} \mathbf{Z}_t \right)^{-1}$.

2 Recursive mean adjustment for pooled estimators

We now turn to the main focus of the paper – that of pooled estimation of the dominant root under homogeneity restrictions.[6] In this section, we assume that the observations

[5] If the assumptions of the factor structure are violated, it is possible that robust nonparametric estimators will dominate the parametric approach. The alternative approaches involve an age-old trade-off to the researcher. The increased demands on long-run covariance parameter estimation of the nonparametric methods may be viewed as the price of flexibility whereas the computational tractability of the parametric method creates the possibility for specification error.

[6] Pooling the cross-section with the time series achieves some bias reduction even for the unadjusted LSDV estimator. Nickell (1981) and Phillips and Sul (2003b) studied the properties of the unadjusted LSDV for the panel AR(1) model when the observations are cross-sectionally independent. Their analyses showed that pooling results in more efficient estimates of ρ, than OLS but does not eliminate the downward bias for finite T even when $N \to \infty$. Because this bias remains when $N \to \infty$ we refer to it as N-asymptotic bias.

are generated as in (3)-(4) above except that $\rho_i = \rho_j = \rho$ for all $i, j = 1, \ldots N$. That is,

$$y_{it} = a_i + \rho y_{it-1} + u_{it}, \tag{10}$$

where $|\rho| \leq 1$, and as before, $u_{it} = \delta_i \theta_t + \epsilon_{it}, \theta_t \overset{iid}{\sim} (0,1), \mathbf{\Sigma} = E[(\epsilon_{1t}, \ldots \epsilon_{Nt})'(\epsilon_{1t}, \ldots \epsilon_{Nt})]$ $=\text{diag}[\sigma_1^2, \ldots \sigma_N^2]$.

Section 2.1 discusses estimation under parameter homogeneity with the RPLS estimator. This is the recursive-mean adjusted analog to the LSDV method pooled estimation with fixed effects [Hsiao (2003)]. While this estimator is not efficient, it provides reduced bias residuals for estimation of error-covariance matrices needed in the construction of the efficient RGLS estimator which we present in subsection 2.2. For expositional clarity, we first develop these estimators under the assumption that the time series follow AR(1) processes. The extension to higher ordered p-th order autoregressions (AR(p)) is taken up in subsection

2.1 Recursive Mean Adjusted Least-Squares Dummy Variable (RLSDV)

RLSDV, introduced here, is the recursive mean adjusted version of the LSDV estimator for regression in panel data with fixed effects. Let $\mu_{it} = (t-1)^{-1} \sum_{s=1}^{t-1} y_{is}$ be the recursively constructed mean. Then the RLSDV estimator is

$$\rho_{\text{RLSDV}} = \frac{\sum_{i=1}^{N} \sum_{t=2}^{T} (y_{it-1} - \mu_{it})(y_{it} - \mu_{it})}{\sum_{i=1}^{N} \sum_{t=2}^{T} (y_{it-1} - \mu_{it})^2}.$$

Analytic expressions for the bias, variance, and mean-square errors are cumbersome and not especially illuminating, so they are relegated to the appendix. Here, we visualize the properties of RLSDV for fixed T as $N \to \infty$ (the $N-$asymptotic properties) in Figure 1 which plots a family of RLSDV $N-$asymptotic bias functions. This is $\underset{N\to\infty}{\text{plim}} \frac{1}{N} (\rho_{\text{RLSDV}} - \rho)$ for alternative sample sizes T, over $0 < \rho < 1$. It can be seen that RLSDV slightly overcompensates for the downward bias and is left with a small upward $N-$asymptotic bias. The resulting upward bias is generally not very sizable and is vanishingly small as $\rho \to 1$.

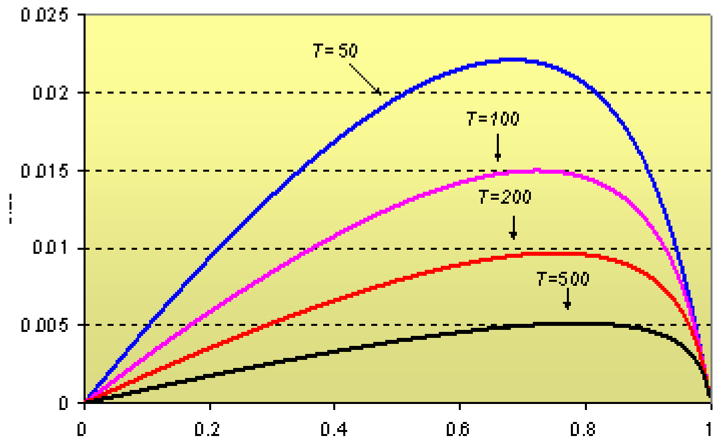

Figure 1: $N-$asymptotic bias of RLSDV for alternative sample sizes T and economically relevant values of ρ

2.2 Pooled recursive mean adjusted GLS with fixed effects (RGLS)

RLSDV does not exploit the cross-sectional covariance structure of the observations in estimation and an efficiency improvement can therefore be achieved by making a mean adjustment to a pooled GLS estimator Since RGLS requires an unbiased estimator of the error-covariance matrix, we employ RLSDV residuals for this purpose. The feasible RGLS estimator is obtained as follows.

Step 1 Estimate ρ by RLSDV as outlined in subsection 2.1.

Step 2 Construct the RLSDV residuals.

$$\hat{u}_{it} = (y_{it} - \rho_{\text{RLSDV}} y_{it-1}) - \frac{1}{T} \sum_{t=1}^{T} (y_{it} - \rho_{\text{RLSDV}} y_{it-1})$$

and use them to estimate the error-covariance matrix $\hat{\mathbf{V}}_u$ as described in step 2 of RSUR. Let v_u^{ij} be the $ij-$th element of $\hat{\mathbf{V}}_u^{-1}$.

Step 3 The RGLS estimator is

$$\underline{\rho}_{\text{RGLS}} = \frac{\sum_{i=1}^{N} \sum_{j=1}^{N} v_u^{ij} \sum_{t=1}^{T} (y_{it} - \mu_{it}) (y_{it-1} - \mu_{it})}{\sum_{i=1}^{N} \sum_{j=1}^{N} v_u^{ij} \sum_{t=1}^{T} (y_{it-1} - \mu_{it})^2}.$$

Characterizing RGLS is analytically cumbersome and provides little insight beyond that discussed in connection with RLSDV. We defer further examination of the small sample properties of RGLS to Section 3 where we compare its performance to alternative estimators in a series of Monte Carlo experiments.

2.3 Extension to Higher-ordered Autoregressions

In the panel AR($p > 1$) case, mean or median unbiased estimation procedures are unavailable due to the difficulty in obtaining analytical formulae for small-sample bias of panel data estimators. A very useful aspect of the recursive mean adjustment procedure is that it can be directly applied to estimate panel AR(p) models. Here, we take an heuristic approach to show how small-sample bias arises and how it can be reduced. We first consider a simple strategy to reduce the bias of autoregressive coefficients under the assumption that the errors are cross-sectionally uncorrelated. We then proceed to implement a GLS version of the estimator under cross section dependence.

Homogeneous coefficients and cross-sectional independence. Let the observations in the panel AR(p) model be generated by

$$\begin{aligned} y_{it} &= \alpha_i + \sum_{j=1}^{p} \delta_j y_{it-j} + u_{it} \\ &= \alpha_i + \rho y_{it-1} + \sum_{j=1}^{p-1} \delta_{j+1} \Delta y_{it-j} + u_{it}. \end{aligned} \tag{11}$$

where $\rho = \sum_{j=1}^{p} \delta_j \;\; u_{it} = \epsilon_{it}$ and $E(u_{it}^2) = \sigma_i^2, E(u_{it}u_{jt}) = 0$ if $i \neq j$ (i.e., the observations are cross-sectionally independent). Let $\underline{z}_{it} = \left(\Delta y'_{it-1}, ...\Delta y'_{it-p+1}\right)'$ and $\underline{\phi}' = (\phi_2, \ldots \phi_p), .$ where $\phi_j = \sum_{i=j+1}^{p} \delta_i$, $j = 1, ..., p-1$. In recursive mean-adjusted form we have

$$(y_{it} - \mu_{it}) = \rho\left(y_{it-1} - \mu_{it}\right) + \underline{z}_{it}\underline{\phi}' + e_{it}. \tag{12}$$

where $\mu_{it} = (t-1)^{-1}\sum_{j=1}^{t-1} y_{jt}$, $e_{it} = \alpha_i - (1-\rho)\mu_{it} + u_{it}$ and $E(e_{it}) = 0$. Estimating $\underline{\beta}' \equiv (\rho, \underline{\phi}')$ by RLSDV as in the AR(1) case substantially reduces bias, resulting in a small positive $N-$asymptotic bias for the estimator of ρ. However, the RLSDV does not adequately address bias in estimating the individual ϕ_j coefficients, which may potentially be even more severe than the bias in unadjusted LSDV. The $N-$asymptotic bias of the estimator for the ϕ_j coefficients is characterized in the appendix. Because half-life estimation in higher-ordered autoregressive systems must be conducted by impulse response analyses, it is important to correct for bias in all of the coefficients.

To reduce bias in estimation of the ϕ_j coefficients, we run a second stage regression by treating the RLSDV estimate as the true value of ρ. Let $y_{it}^{\dagger} \equiv y_{it} - \rho_{\mathrm{RLSDV}} y_{it-1}$ and run the regression

$$y_{it}^{\dagger} = \underline{z}_{it}\underline{\phi}' + \tilde{u}_{it} \tag{13}$$

by OLS. Call the resulting estimator $\underline{\hat{\phi}}^{\dagger}$. We show below that the residual bias in $\hat{\phi}^{\dagger}$ is inconsequential.

Heterogeneous coefficients and cross-sectional independence To extend this bias reduction method to heterogeneous AR(p) coefficients, we suggest the following procedure.

Step 1 Regress $(y_{it} - \mu_{it})$ on $\underline{z}_{it}$ without a constant. Denote the regression residuals by $\hat{v}_{it}$. Also, regress $(y_{it-1} - \mu_{it})$ on $\underline{z}_{it}$ without a constant and denote these regression residuals by $\hat{v}_{it-1}$. The ρ_{LSDV} is obtained by running the pooled regression of $\hat{v}_{it}$ on $\hat{v}_{it-1}$ without a constant.

Step 2 Let $y_{it}^{\dagger} \equiv y_{it} - \rho_{\text{LSDV}} y_{it-1}$. Regressing $y_{it}^{\dagger}$ on $\underline{z}_{it}$ with a constant term for each i gives the estimator $\underline{\hat{\phi}}_i^{\dagger}$. Standard errors for ρ_{LSDV} and $\underline{\hat{\phi}}_i^{\dagger}$ must be constructed using $u_{it}^{\dagger} = \hat{u}_{it} - \frac{1}{T}\sum_{t=1}^{T}\hat{u}_{it}$ where $\widehat{u}_{it} = y_{it} - \rho_{\text{LSDV}} y_{it-1} - \underline{z}_{it}'\underline{\hat{\phi}}_i^{\dagger}$.

Heterogeneous coefficients and cross-sectional dependence Under cross section dependence we extend the above procedure to an SUR estimator as follows.

Step 1 Obtain $u_{it}^{\dagger}$ as described above. Estimate the covariance matrix $\hat{V}_u$ as described in steps 1-2 in subsection 1.2. Let v_u^{ij} be the $ij-$th element of $\hat{V}_u^{-1}$.

Step 2 Let $\mathbf{Z}_{t-1} = \text{diag}\left[\underline{z}_{1t}, \underline{z}_{2t}, \ldots \underline{z}_{Nt}\right]$, $\underline{Y}_t = (y_{1t} - \mu_{1t}, \ \ldots \ , y_{Nt} - \mu_{Nt})'$ and $\underline{Y}_{t-1} = \left(y_{1t-1} - \mu_{1t-1}, \ \ldots \ , y_{Nt-1} - \mu_{Nt-1}\right)'$, $\hat{\mathbf{w}}_t = \underline{Y}_t - \left(\sum_{t=1}^{T}\mathbf{Z}_{t-1}'\hat{\mathbf{V}}_u^{-1}\mathbf{Z}_{t-1}\right)^{-1}\sum_{t=1}^{T}\mathbf{Z}_{t-1}'\hat{\mathbf{V}}_u^{-1}\underline{Y}_t$, and $\hat{\mathbf{w}}_{t-1} = \underline{Y}_{t-1} - \left(\sum_{t=1}^{T}\mathbf{Z}_{t-1}'\hat{\mathbf{V}}_u^{-1}\mathbf{Z}_{t-1}\right)^{-1}\sum_{t=1}^{T}\mathbf{Z}_{t-1}'\hat{\mathbf{V}}_u^{-1}\underline{Y}_{t-1}$. Let $\underline{\beta} = (\rho_1, \ldots \rho_N)'$, then the RSUR estimator for β is

$$\beta_{\text{RSUR}} = \left(\sum_{t=1}^{T}\hat{\mathbf{w}}_{t-1}'\hat{\mathbf{V}}_u^{-1}\hat{\mathbf{w}}_{t-1}\right)^{-1}\sum_{t=1}^{T}\hat{\mathbf{w}}_{t-1}'\hat{\mathbf{V}}_u^{-1}\mathbf{w}_t$$

Step 3 Let $\tilde{y}_{it}^{\dagger} = \tilde{y}_{it} - \beta_{\text{RSUR}}\tilde{y}_{it-1}$, and $\underline{\tilde{Y}}_t^{\dagger} = \left(\tilde{y}_{1t}^{\dagger}, \ldots\tilde{y}_{NT}^{\dagger}\right)$. The RSUR estimator for $\underline{\phi}$ is

$$\phi_{\text{RSUR}} = \left(\sum_{t=1}^{T}\mathbf{Z}_{t-1}'\hat{\mathbf{V}}_u^{-1}\mathbf{Z}_{t-1}\right)^{-1}\sum_{t=1}^{T}\mathbf{Z}_{t-1}'\hat{\mathbf{V}}_u^{-1}\underline{\tilde{Y}}_t^{\dagger}$$

3 Monte Carlo Experiments

In this section, we examine the small sample performance of the alternative estimators. Subsection 3.1 reports performance results for the RSUR estimator and subsection 3.2 reports results for pooled estimators. The observations are generated with a homogeneous autoregressive coefficient and without a constant. The data generating process is

$$
\begin{aligned}
y_{it} &= \rho y_{it-1} + u_{it} \\
u_{it} &= \delta_i \theta_t + e_{it} \\
E(u_{it} u_{jt}) &= \begin{cases} \delta_i \delta_j \text{ if } i \neq j \\ \delta_i^2 + \sigma_i^2 \text{ if } i = j. \end{cases}
\end{aligned}
$$

The parameters of the data generating process are set in line with estimates that are typical in macroeconomic empirical studies. We consider combinations of $N \in \{5, 10, 20\}$ and $T \in \{50, 100, 200\}$ for $\rho \in \{0.9, 0.95, 0.99\}$ unless specified otherwise. δ_i is randomly generated on U[1,4] where U is the uniform distribution. The resulting cross sectional correlations lie in the interval $(0.5, 0.99)$. Each simulation experiment is based on 10,000 replications.

3.1 Small Sample Performance of RSUR

We first report results for nonparametric and parametric RSUR in which heterogenous constants and autocorrelation coefficients were allowed in estimation. To evaluate the value-added of RSUR, we compare it against OLS as a benchmark estimator. The bias and mean squared error for RSUR presented in Table 1 refers to the cross-sectional average, $N^{-1} \sum_{i=1}^{N} \rho_{\text{RSUR},i}$. From the table, it can be seen that RSUR achieves substantial bias reduction over OLS but a small amount of downward bias. RSUR is relatively insensitive to the number of cross-sectional units N. The performance of the parametric RSUR is superior to nonparametric RSUR in that it exhibits less bias and has lower mean squared error.

Table 2 reports the effective size of the Wald test of the homogeneity restrictions on the autocorrelation coefficients. Again, parametric RSUR exhibits better performance. While the test constructed using nonparametric RSUR becomes badly oversized for the large systems $(N = 20)$, the performance for parametric RSUR remains accurately sized. Moreover, the effective size of this test is not sensitive either to the size of the time-series T or to the size of the cross-section N.

3.2 Small Sample Performance of Pooled Estimator

If the data pass the homogeneity test, estimation proceeds with a pooled estimator. We now undertake an evaluation of the performance of our RGLS estimator. In the next section, we compare performance of the recursive mean adjusted estimators to panel mean-unbiased estimators suggested by Phillips and Sul (2003a,b).[7] The idea behind mean-unbiased estimators to estimate ρ using a method with known bias, such as LSDV and then to adjust the point estimate by this bias. Specifically we discuss mean adjusted LSDV (MLSDV) and mean adjusted GLS (MGLS).[8] Since these estimators are described in detail in Phillips and Sul (2003 a, b), we provide only a cursory description here.

MSLDV proceeds as follows: Nickell (1981) derived the bias in LSDV, $\text{plim}_{N\to\infty}(\rho_{\text{LSDV}}) = m(\rho)$ (with the precise formula for the function $m(\cdot)$ given in the appendix). Then the MLSDV estimator is obtained by inverting this function, $\rho_{\text{MSLDV}} = m^{-1}(\rho_{\text{LSDV}})$. For MGLS, we draw on Phillips and Sul (2003a) who show that in an environment where the regression errors are cross-sectionally correlated and follow a single factor structure, the $N-$asymptotic bias of the pooled GLS estimator depends neither on the parameters δ_i nor on the unobserved factor θ_t. As a result of this independence we can apply the same mean adjustment rule to get MGLS, namely $\rho_{\text{MGLS}} = m^{-1}(\rho_{\text{GLS}})$.

We compare its performance to the following alternative pooled estimators: LSDV, MLSDV, nonparametric MGLS and parametric MGLS. The LSDV (and MLSDV) estimators include a set of time-specific dummy variables as a partial control for cross-sectional dependence.

The results, reported in Table 3, list the alternative estimators in inverse order of their bias. RGLS is the least biased. Nonparametric MGLS is less biased than MLSDV but the GLS estimator is more efficient as expected while parametric MGLS achieves additional efficiency gains. RGLS is the preferred estimator as it generally dominates the others in reducing bias and is most efficient in a mean-square error sense. The exception is for $T = 200$ where RGLS appears to slightly overcorrect the bias (e.g., for $N = 20$, $T = 200$, $\rho = 0.95$, the bias is 0.00235).

A visual representation of the alternative estimator performance is given in Figure 2 which displays empirical distributions for each of the estimators of $\rho = 0.98$ with $N = 21, T = 51$. Here again, it can be seen that the worst performer is pooled OLS which is has large variance and is severely downward biased. Next, is LSDV, followed by nonparametric MGLS, parametric MGLS and our preferred RGLS estimator.

[7] In the AR(p) model with exogenous regressors under cross section independence, IV/GMM estimators in dynamic panel regressions can be employed to reduce bias but neither the asymptotic nor the finite sample properties of these estimators have been investigated under cross-sectional dependence. See Ahn and Schmidt (1995) and Hahn and Kuersteiner (2001).

[8] In this paper we consider only mean-unbiased estimators since for moderately sized T samples, the properties of median-unbiased and mean-unbiased estimators are quite similar.

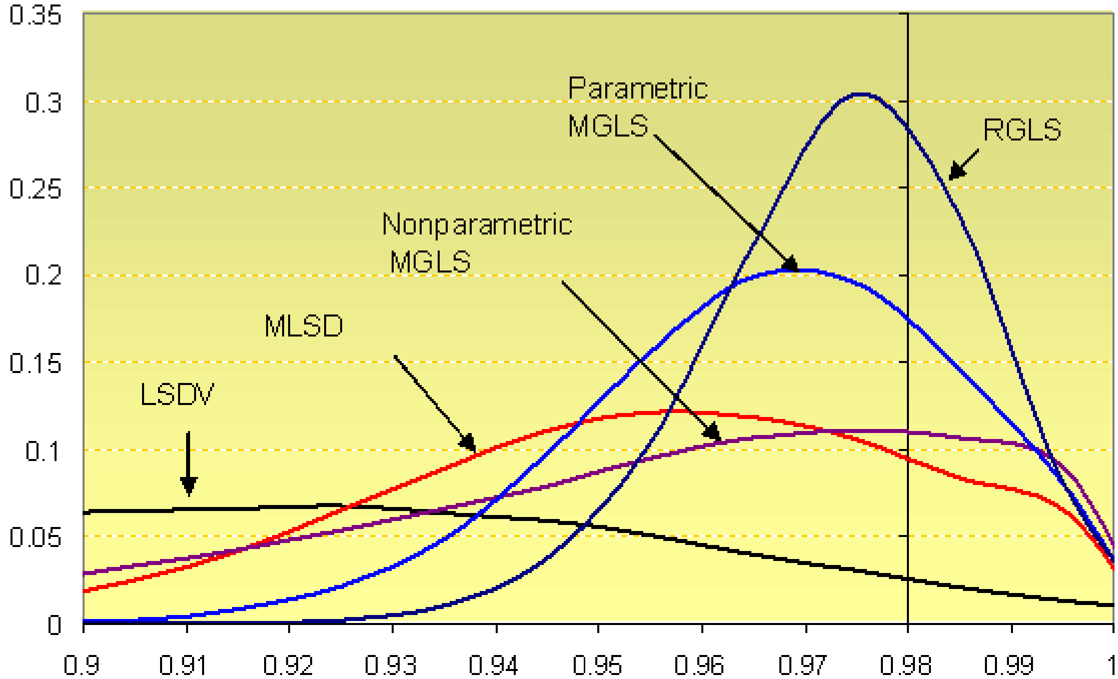

Figure 2: Empirical distributions of alternative pooled estimators. $\rho = 0.98$, N=21, T=51.

Table 4 shows the results for the AR(2) case. The DGP is

$$\begin{aligned} y_{it} &= \rho y_{it-1} + e_{it} \\ e_{it} &= \phi_i e_{it-1} + u_{it} \\ u_{it} &= \delta_i \theta_t + \varepsilon_{it} \end{aligned}$$

where $\varepsilon_{it} \overset{iid}{\sim} N(0,1)$ and $\phi_i \sim U[-0.3,-0.2]$ if $\psi = -0.3$ and $\phi_i \sim U[0.3,0.4]$ if $\psi = 0.3$. As can be seen, RGLS is very efficient in terms of relative mean square error. For small $T = 50$ and large $N = 50$, RGLS retains a slight negative bias. As T increases and N decreases, RGLS begins to overcompensate slightly as ρ moves down away from 1, however, the overall amount of residual bias in RGLS is very small.

4 Estimating Short-Term Interest Rate Dynamics

This section uses RGLS to estimate dynamic models of short-term interest rates. Much interest has centered on mean-reverting specifications suggested by Cox, Ingersoll, and Ross (1985, hereafter CIR). Based on a single factor model of the term structure CIR characterize the interest rate as the continuous time first-order autoregressive process

$$dr = \kappa(\theta - r)dt + \sigma\sqrt{r}dZ, \tag{14}$$

where θ represents long-term mean of the interest rate to which r reverts at rate κ and dZ is a standard Wiener process.[9] In practice the parameters in (14) are estimated using a single time series of short-term interest rates which itself serves as a proxy for the instantaneous interest rate. Models of the term structure of interest rates are often evaluated on the basis of those estimates using available closed-form expressions for default-free bond prices. As pointed out by Ball and Torous (1996), the downward bias in these estimates produce implied adjustment speeds of the short term interest rate that are too high to be consistent with the actual dynamics of bond price data.

To obtain reduced biased estimates of short-term interest dynamics, we follow Ball and Torus in discretizing (14) as

$$r_{t+\Delta} = \alpha + \beta r_t + u_t, \tag{15}$$

where $\alpha = \theta(1 - e^{-\kappa\Delta t})$, $\beta = e^{-\kappa\Delta}$ and $u_t = \sigma\epsilon_t\sqrt{\Delta}$ where ϵ_t is zero meaned with variance proportional to r_t.

4.1 Another Simulation Experiment

Before we employ the RGLS estimator on the data, we examine its performance in the context of the simulation experiment considered by Ball and Torous (1996) which we extend to the panel context. We consider two alternative cross section dimensions of $N = 5, 10$ which are comparable. Pseudo observations for annualized interest rates are generated from a first order discrete-time approximation to (14) assuming 360 time steps per year. We consider weekly frequencies ($\Delta = 1/52$) and monthly frequencies ($\Delta = 1/12$) for sampling the data. and data sets consisting of five years and ten years of observations. Parameter values are set according to $\kappa = 1$, and $\theta \sim U[0.1, 0.9]$. The initial observation for interest rate is set at θ, and $\sigma = 0.05$ and 0.1, respectively. Cross-sectional dependence is introduced by setting $\text{cov}(dZ_i, dZ_j) = \rho_{ij}dt$ where $\rho_{ij} \sim U[-0.5, 0.5]$. Each simulation run is carried out with 1,000 as in Ball and Torous (1996).

Table 4 reports the performance of LSDV and RGLS with associated biases $(\hat{\kappa} - \kappa)$ in parentheses. There is only a tiny bit of residual bias in the RGLS estimator. Better results are evidently obtained when observations are sampled weekly.

[9] Chan et. al. (1992) examined models of the form $dr = (\alpha + \beta r)\, dt + \sigma r^{\gamma} dZ$. The specification that we estimate is a special case of Chan et. al's. that is obtained with $\kappa = 1$, $\alpha = \theta$, $\beta = -1$ and $\gamma = 0.5$. This characterization guarantees that the interest rate is nonnegative.

4.2 Five Weekly Eurocurrency Rates

We now apply RGLS to estimate the speed of adjustment in a data set of weekly eurocurrency rates.[10] Eurocurrency rates are free market interest rates which are not directly affected by such domestic factors such as reserve requirements and foreign exchange controls. We collected five weekly data sets of 325 weekly Eurodeposit rates for Australia, Canada, Germany, Switzerland, and the U.S.–one data set for each day of the week. The time span of the data extends from December 1, 1997 to February 20, 2004. These data were originally collected by Harris bank which we downloaded from *Datastream.* All rates are expressed in percent per annum

In our empirical specification, we allow the interest rates to follow an AR(2). The Table 5 reports the results for ρ estimated for each of the five data sets. We first estimated the system by RSUR and conduct tests of homogeneity restrictions across the five eurocurrency rates. None of the p-values for this test lie below 0.3 so proceed to perform estimation under homogeneity restrictions. Panel estimation by LSDV yields estimates that consistently lie below 1 whereas the RGLS estimates consistently lie above 1. The bias corrected RGLS estimator suggests that short-term interest rates do not mean revert and instead may be unit-root nonstationary. However, a 95% confidence interval includes values of ρ that lie below 1.It is also possible that short-term interest rates are governed by stationary but nonlinear dynamics.

5 Conclusion

In this paper, we extend the idea of recursive mean adjustment as a bias reduction strategy in estimating the dominant root in dynamic systems and for panel regressions. The method is straightforward to implement and has familiar asymptotic properties. Because mean and median unbiased estimators are generally unavailable for higher ordered panel autoregression models, the recursive mean adjustment procedure fills an important gap in the literature. Monte Carlo experiments show that it is an efficient and effective bias reduction strategy in small samples.

We applied RGLS to estimate the speed of adjustment in a small panel of short-term interest rates. The bias-adjusted estimates suggest that there may be no reversion to the mean.

[10] Factor models of the term structure are concerned with mean reversion in real interest rates but due to the very short sampling intervals, our examination focuses on the nominal interest rate.

References

[1] Ahn, S.C., Schmidt P., 1995. Efficient Estimation of Models for Dynamic Panel-data. *Journal of Econometrics* 68, 5–27.

[2] Andrews, D.W.K., 1993. Exactly Median-unbiased Estimation of First Order Autoregressive/Unit root Models. *Econometrica* 61, 139–165.

[3] Andrews, D.W.K., Chen, H., 1994. Approximately median-unbiased estimation of autoregressive models. *Journal of Business and Economic Statistics* 12, 187–204.

[4] Ball, Clifford A. and Walter N. Torous, 1996. "Unit Roots and the Estimation of Interest Rate Dynamics," *Journal of Empirical Finance*, 3, 25-238.

[5] Bai, J., Ng, S., 2002. Determining the Number of Factors in Approximate Factor Models. *Econometrica* 70, 191–221.

[6] Chan, K., G. Karolyi„ F. Longstaff, and A. Sanders,1992. "An Empirical Comparison of Alternative Models of the Short-Teerm Interest Rate," *Journal of Finance*, 47, 1209-27.

[7] Frankel, Jeffrey A. and Andrew K. Rose, 1996. "A Panel Project on Purchasing Power Parity: Mean Reversion within and between Countries," *Journal of International Economics,* 40, 209-24

[8] Hahn J, Kuersteiner G., 2002. Asymptotically Unbiased Inference for a Dynamic Panel Model with Fixed Effects When Both n and T are Large. *Econometrica* 70(4), 1639–1657.

[9] Gibbons, M.R. and K. Ramaswamy, 1993."A Test of the Cox, Ingersoll, and Ross Model of the Term Structure," *The Review of Financial Studies*, 6, 619—58.

[10] Hsiao, C., 2003. Analysis of Panel Data, Second Edition. Cambridge University Press.

[11] Kendall, M.G., 1954. Note on Bias in the Estimation of Autocorrelation. *Biometrika* 41, 403–404.

[12] Marriott, F.H.C., Pope, J.A., 1954. Bias in the estimation of autocorrelations, *Biometrika* 41, 393–403.

[13] Moon, H.R., Perron, B., 2003. Testing for A Unit Root in Panels with Dynamic Factors. *Journal of Econometrics, forthcoming.*

[14] Murray, C.J., Papell, D.H., 2002. The Purchasing Power Parity Persistence Paradigm. *Journal of International Economics* 56(1), 1–19.

[15] Nickell, S. 1981. Biases in Dynamic Models with Fixed Effects. *Econometrica* 49, 1417–1426.

[16] Papell, David H., 2004. "The Panel Purchasing Power Parity Puzzle", *mimeo*, University of Houston.

[17] Perarson, N.D. and T. Sun, 1994. "Exploiting the Conditional Density in Estimating the Term STructure: An application to the Cox, Ingersoll, and Ross Model," *Journal of Finance*, 49, 1279–1304.

[18] Phillips, P.C.B. and D. Sul, 2003a. "Dynamic Panel Estimation and Homogeneity Testing Under Cross Section Dependence," *The Econometrics Journal* 6, 217—-259.

[19] Phillips, P.C.B. and D. Sul, 2003b. "Bias in Dynamic Panel Estimation with Fixed Effects, Incidental Trends and Cross Section Dependence," *mimeo*, University of Auckland.

[20] So, B.S., Shin, D.W., 1999. Recursive Mean Adjustment in Time Series Inferences. *Statistics and Probability Letters* 43, 65–73.

[21] Wu, Yangru and Hua Zhang, 1996. "Mean Reversion in Interest Rates: New Evidence from a Panel of OECD Countries," *Journal of Money, Credit and Banking*, 28, 604-621.

Table 1: Monte Carlo Performance of RSUR

			Bias × T			Relative MSE[a]	
N	T	ρ	OLS	RSUR[b]	RSUR[c]	RSUR[b]	RSUR[c]
		0.900	-4.350	-1.343	-1.248	0.216	0.204
5	50	0.950	-4.880	-1.692	-1.604	0.200	0.189
		0.990	-5.234	-1.686	-1.620	0.168	0.160
		0.900	-4.071	-0.780	-0.729	0.214	0.210
5	100	0.950	-4.608	-1.317	-1.261	0.191	0.185
		0.990	-5.322	-1.678	-1.635	0.158	0.154
		0.900	-3.987	-0.204	-0.182	0.228	0.227
5	200	0.950	-4.417	-0.793	-0.764	0.192	0.190
		0.990	-5.350	-1.644	-1.617	0.157	0.154
		0.900	-4.414	-1.501	-1.225	0.205	0.167
10	50	0.950	-4.909	-1.830	-1.579	0.195	0.161
		0.990	-5.251	-1.787	-1.607	0.165	0.140
		0.900	-4.186	-0.879	-0.717	0.176	0.161
10	100	0.950	-4.688	-1.422	-1.238	0.167	0.149
		0.990	-5.388	-1.729	-1.603	0.144	0.131
		0.900	-3.965	-0.198	-0.123	0.182	0.177
10	200	0.950	-4.373	-0.796	-0.698	0.156	0.150
		0.990	-5.323	-1.645	-1.557	0.138	0.131
		0.900	-4.387	-1.617	-1.212	0.254	0.168
20	50	0.950	-4.913	-1.891	-1.564	0.228	0.161
		0.990	-5.245	-1.791	-1.594	0.183	0.141
		0.900	-4.144	-1.016	-0.702	0.202	0.163
20	100	0.950	-4.653	-1.584	-1.231	0.194	0.153
		0.990	-5.339	-1.838	-1.622	0.163	0.136
		0.900	-4.002	-0.305	-0.155	0.190	0.177
20	200	0.950	-4.401	-0.927	-0.725	0.168	0.152
		0.990	-5.307	-1.747	-1.574	0.152	0.135

Notes: [a]MSE of RSUR relative to MSE of OLS. [b]Nonparametric RSUR. [c]Parametric RSUR.

Table 2: Monte Carlo Performance of Test for Homogeneity

			RSUR[a/]			RSUR[b/]		
N	T	ρ	10%	5%	1%	10%	5%	1%
		0.900	0.193	0.118	0.040	0.161	0.092	0.027
5	50	0.950	0.196	0.120	0.040	0.166	0.097	0.028
		0.990	0.176	0.107	0.033	0.150	0.088	0.024
		0.900	0.150	0.087	0.025	0.140	0.079	0.020
5	100	0.950	0.165	0.097	0.028	0.154	0.087	0.022
		0.990	0.154	0.088	0.024	0.144	0.080	0.020
		0.900	0.123	0.066	0.017	0.116	0.062	0.015
5	200	0.950	0.133	0.074	0.018	0.127	0.069	0.016
		0.990	0.156	0.085	0.020	0.149	0.081	0.020
		0.900	0.354	0.253	0.114	0.194	0.118	0.035
10	50	0.950	0.354	0.253	0.117	0.204	0.123	0.038
		0.990	0.303	0.209	0.090	0.180	0.106	0.030
		0.900	0.229	0.140	0.048	0.156	0.089	0.022
10	100	0.950	0.254	0.160	0.055	0.182	0.105	0.029
		0.990	0.238	0.148	0.048	0.179	0.102	0.027
		0.900	0.158	0.089	0.021	0.124	0.067	0.014
10	200	0.950	0.180	0.104	0.028	0.149	0.082	0.019
		0.990	0.204	0.122	0.040	0.178	0.103	0.029
		0.900	0.780	0.703	0.540	0.235	0.144	0.045
20	50	0.950	0.746	0.669	0.511	0.249	0.151	0.048
		0.990	0.631	0.550	0.398	0.219	0.131	0.040
		0.900	0.453	0.344	0.170	0.175	0.102	0.027
20	100	0.950	0.495	0.380	0.202	0.216	0.128	0.040
		0.990	0.443	0.330	0.164	0.217	0.128	0.037
		0.900	0.263	0.171	0.060	0.149	0.079	0.019
20	200	0.950	0.299	0.201	0.076	0.177	0.099	0.029
		0.990	0.334	0.226	0.088	0.216	0.130	0.041

Notes: [a/]Nonparametric RSUR. [b/]Parametric RSUR.

Table 3: Monte Carlo Performance of Pooled Estimators

			Bias × T					Relative MSE			
N	T	ρ	A	B	C	D	E	B/A	C/A	D/A	E/A
		0.90	-3.975	-0.960	-0.813	-0.715	-0.283	0.280	0.172	0.161	0.116
5	50	0.95	-4.456	-1.274	-1.081	-0.973	-0.618	0.246	0.163	0.150	0.093
		0.99	-4.745	-1.552	-1.248	-1.162	-0.605	0.203	0.137	0.124	0.069
		0.90	-3.710	-0.814	-0.601	-0.546	0.188	0.310	0.177	0.171	0.144
5	100	0.95	-4.188	-1.022	-0.791	-0.729	-0.297	0.268	0.157	0.151	0.105
		0.99	-4.826	-1.508	-1.174	-1.121	-0.680	0.208	0.131	0.125	0.069
		0.90	-3.679	-0.828	-0.496	-0.469	0.692	0.347	0.186	0.183	0.177
5	200	0.95	-4.048	-0.952	-0.616	-0.583	0.168	0.295	0.161	0.158	0.130
		0.99	-4.859	-1.423	-1.049	-1.012	-0.670	0.223	0.132	0.128	0.076
		0.90	-3.975	-0.743	-0.852	-0.529	-0.109	0.189	0.124	0.081	0.056
10	50	0.95	-4.412	-0.991	-1.101	-0.747	-0.433	0.164	0.130	0.081	0.045
		0.99	-4.679	-1.199	-1.102	-0.805	-0.409	0.125	0.109	0.066	0.032
		0.90	-3.769	-0.664	-0.567	-0.382	0.348	0.221	0.105	0.087	0.075
10	100	0.95	-4.198	-0.819	-0.748	-0.535	-0.114	0.185	0.097	0.077	0.051
		0.99	-4.815	-1.241	-1.022	-0.822	-0.492	0.136	0.087	0.067	0.033
		0.90	-3.604	-0.641	-0.369	-0.276	0.882	0.259	0.100	0.094	0.101
10	200	0.95	-3.938	-0.723	-0.466	-0.360	0.382	0.212	0.086	0.079	0.069
		0.99	-4.764	-1.157	-0.884	-0.761	-0.468	0.151	0.078	0.069	0.037
		0.90	-3.861	-0.667	-1.052	-0.414	0.009	0.174	0.180	0.047	0.031
20	50	0.95	-4.323	-0.914	-1.276	-0.627	-0.317	0.146	0.186	0.048	0.024
		0.99	-4.566	-1.091	-1.100	-0.594	-0.287	0.107	0.144	0.039	0.017
		0.90	-3.648	-0.589	-0.665	-0.276	0.455	0.198	0.096	0.047	0.045
20	100	0.95	-4.065	-0.732	-0.856	-0.418	-0.001	0.164	0.094	0.042	0.027
		0.99	-4.648	-1.110	-1.032	-0.657	-0.383	0.117	0.091	0.042	0.018
		0.90	-3.555	-0.590	-0.407	-0.203	0.954	0.235	0.066	0.051	0.065
20	200	0.95	-3.869	-0.663	-0.500	-0.268	0.470	0.190	0.059	0.043	0.041
		0.99	-4.629	-1.054	-0.886	-0.635	-0.366	0.131	0.062	0.041	0.020

(A): LSDV; (B): MLSDV; (C): Nonparametric MGLS; (D) Parametric MGLS; (E): Parametric RGLS.

Table 4: Finite Sample Performance of RGLS for AR(2) case

				$\hat{\rho}$					$\hat{\phi}$				
				Bias × T			MSE Ratio		Bias × T			MSE Ratio	
T	N	ρ	ψ	A	B	C	B/A	C/A	D	E	F	E/D	F/D
		0.90	0.3	-2.487	0.386	1.107	0.902	0.260	147.921	0.721	1.110	0.021	0.008
			-0.3	-4.401	0.251	1.526	0.827	0.240	-114.999	-0.479	0.509	0.035	0.014
		0.92	0.3	-2.461	0.290	1.006	0.877	0.254	153.343	0.725	1.102	0.019	0.008
			-0.3	-4.449	0.044	1.332	0.791	0.228	-116.994	-0.570	0.426	0.034	0.013
500	5	0.95	0.3	-2.433	0.070	0.780	0.814	0.234	161.475	0.705	1.067	0.017	0.007
			-0.3	-4.562	-0.405	0.899	0.707	0.197	-119.986	-0.769	0.239	0.032	0.013
		0.98	0.3	-2.474	-0.373	0.326	0.656	0.176	169.606	0.633	0.980	0.015	0.006
			-0.3	-4.912	-1.243	0.048	0.541	0.134	-122.978	-1.121	-0.119	0.031	0.012
		1.00	0.3	-2.895	-0.826	-0.263	0.275	0.064	175.028	2.482	2.721	0.015	0.006
			-0.3	-5.569	-1.594	-0.508	0.275	0.064	-124.972	-0.736	0.136	0.030	0.012
		0.90	0.3	-2.609	-0.038	0.963	0.832	0.120	72.558	0.512	0.941	0.044	0.011
			-0.3	-4.682	-0.540	1.207	0.730	0.100	-59.232	-0.788	0.354	0.068	0.016
		0.92	0.3	-2.610	-0.152	0.845	0.795	0.113	75.179	0.500	0.917	0.041	0.010
			-0.3	-4.781	-0.768	0.993	0.687	0.090	-60.253	-0.890	0.259	0.066	0.016
250	10	0.95	0.3	-2.635	-0.392	0.594	0.712	0.094	79.110	0.467	0.863	0.036	0.009
			-0.3	-5.016	-1.236	0.538	0.599	0.069	-61.785	-1.094	0.062	0.062	0.015
		0.98	0.3	-2.801	-0.834	0.127	0.551	0.056	83.041	0.437	0.799	0.033	0.008
			-0.3	-5.595	-1.966	-0.244	0.464	0.043	-63.317	-1.347	-0.231	0.059	0.014
		1.00	0.3	-3.095	-0.959	-0.127	0.305	0.024	85.661	2.293	2.534	0.031	0.008
			-0.3	-5.873	-1.831	-0.240	0.304	0.024	-64.339	-0.812	0.207	0.057	0.014

Table 4: Continued–

T	N	ρ	ψ	$\hat{\rho}$ Bias × T A	B	C	MSE Ratio B/A	C/A	$\hat{\phi}$ Bias × T D	E	F	MSE Ratio E/D	F/D
		0.90	0.3	-2.703	-0.527	0.592	0.680	0.049	29.680	0.340	0.796	0.104	0.022
			-0.3	-4.970	-1.366	0.549	0.573	0.031	-22.882	-1.164	0.059	0.183	0.036
		0.92	0.3	-2.725	-0.634	0.466	0.637	0.041	30.771	0.351	0.783	0.096	0.021
			-0.3	-5.127	-1.568	0.337	0.537	0.025	-23.279	-1.236	-0.025	0.178	0.035
100	25	0.95	0.3	-2.808	-0.848	0.214	0.557	0.027	32.406	0.395	0.772	0.086	0.019
			-0.3	-5.463	-1.923	-0.062	0.473	0.018	-23.875	-1.333	-0.162	0.169	0.034
		0.98	0.3	-3.054	-1.139	-0.164	0.442	0.017	34.041	0.671	0.931	0.078	0.017
			-0.3	-5.944	-2.217	-0.455	0.388	0.016	-24.471	-1.285	-0.188	0.161	0.032
		1.00	0.3	-3.035	-0.942	-0.053	0.310	0.010	35.132	2.386	2.576	0.078	0.021
			-0.3	-5.781	-1.823	-0.087	0.310	0.010	-24.868	-0.782	0.278	0.154	0.031
		0.90	0.3	-2.805	-0.769	0.270	0.575	0.019	14.944	0.374	0.786	0.213	0.051
			-0.3	-5.158	-1.669	0.056	0.485	0.012	-11.313	-1.138	-0.126	0.392	0.083
		0.92	0.3	-2.852	-0.866	0.146	0.537	0.015	15.496	0.421	0.808	0.198	0.047
			-0.3	-5.339	-1.825	-0.111	0.457	0.011	-11.510	-1.171	-0.169	0.380	0.081
50	50	0.95	0.3	-2.979	-1.036	-0.076	0.473	0.011	16.323	0.569	0.908	0.179	0.043
			-0.3	-5.647	-2.036	-0.339	0.410	0.011	-11.805	-1.173	-0.188	0.362	0.078
		0.98	0.3	-3.171	-1.159	-0.253	0.386	0.012	17.151	1.080	1.369	0.166	0.043
			-0.3	-5.845	-2.044	-0.347	0.351	0.009	-12.100	-0.996	-0.023	0.342	0.074
		1.00	0.3	-2.987	-0.876	-0.005	0.306	0.006	17.703	2.384	2.676	0.171	0.057
			-0.3	-5.614	-1.695	0.004	0.303	0.006	-12.297	-0.640	0.315	0.326	0.073

(A) LSDV for ρ; (B) RLSDV estimator for ρ; (C) RGLS estimator for ρ; (D) LSDV for ϕ_i; (E) RLSDV estimator for ϕ_i; (F) RGLS estimator for ϕ_i

Table 5: Simulation Results for Short-Term Interest Rate Dynamics $\kappa = 1$, $\sigma = 0.05$ .Bias $(\hat{\kappa} - \kappa)$ in parentheses.

Time Span	N=5		N=10	
	LSDV	PRGLS	LSDV	PRGLS
60 months	0.8676	0.9185	0.8739	0.9260
	(0.59)	(-0.02)	(0.51)	(-0.11)
120 months	0.8974	0.9239	0.8990	0.9259
	(0.23)	(-0.09)	(0.21)	(-0.11)
260 weeks	0.9692	0.9809	0.9699	0.9819
	(0.60)	(-0.01)	(0.56)	(-0.05)
520 weeks	0.9756	0.9818	0.9759	0.9822
	(0.27)	(-0.05)	(0.25)	(-0.07)

Table 6: Estimates for the Dynamics of Eurocurrency Rates

		Mon	Tue	Wed	Thr	Fri
Homogeneity Test	p-value	(0.5008)	(0.4983)	(0.3373)	(0.3515)	(0.4250)
LSDV	ρ	0.9895	0.9930	0.9958	0.9958	0.9892
RGLS	ρ	1.0002	1.0014	1.0023	1.0025	1.0004
	$[\rho_{5\%}]$	[0.9928]	[0.9940]	[0.9958]	[0.9936]	[0.9933]

Notes: 5% denotes the lower end of 95% confidence inteval of RGLS estimates.

Appendix

N-asymptotic bias of LSDV

Nickell (1981) shows that the bias in the LSDV estimator for the AR(1) model with fixed effects is

$$\begin{aligned} \operatorname*{plim}_{N\to\infty}(\rho_{\mathrm{LSDV}}) \equiv\; & m(\rho) = \rho - \left(\frac{1+\rho}{T-1}\right)\left[1 - \frac{1}{T}\frac{1-\rho^T}{1-\rho}\right] \\ & \times \left\{1 - \left(\frac{1}{T-1}\right)\left(\frac{2\rho}{1-\rho}\right)\left[1 - \frac{1}{T}\frac{1-\rho^T}{1-\rho}\right]\right\}^{-1}, \end{aligned} \tag{16}$$

N-Asymptotic bias, variance, and mean-square error of RLSDV

Here, we derive the bias, variance, and mean-square errors that underlies Figure 1. Before proceeding, we require some preliminary results. First, we adopt

Assumption A1: (Regularity conditions on the error term) The ε_{it} have zero mean, finite $2+2\nu$ moments for some $\nu > 0$, are independent over i and t with $E(\varepsilon_{it}^2) = \sigma_i^2$ for all t, and $\lim_{N\to\infty} \frac{1}{N}\sum_{i=1}^N \sigma_i^2 = \sigma^2$.

We also make use of

Lemma 1 $E\left(\sum_{s=1}^{t-1} x_{is}\right)^2 = \left(\frac{1-\rho^2}{(1-\rho)^2}\right)\left[t-1-2\rho\frac{1-\rho^{t-1}}{1-\rho^2}\right]\sigma_x^2.$

Lemma 2 $E\left(x_{it-1}\sum_{s=1}^{t-1} x_{is}\right) = \sigma_x^2\left(\frac{1-\rho^{t-1}}{1-\rho}\right).$

Proofs to lemmas 1 and 2 are straightforward and are omitted.

Lemma 3 *(Asymptotic Equivalence)* $T^{-1}\sum_{t=1}^T Ee_{it}e_{it}' - T^{-1}\sum_{t=1}^T Eu_{it}u_{it}' = O\left(T^{-1}\ln T\right)$

Proof. Let $A_i = diag\left(\rho_i\right)$, then we have

$$\begin{aligned} T^{-1}\sum_{t=2}^T Ee_{it}e_{it}' \;=\;& T^{-1}\sum_{t=2}^T (I-A_i)\left(\frac{1}{(t-1)^2}\right)E\left[\sum_{s=1}^{t-1} x_{is}\right]'\left[\sum_{s=1}^{t-1} x_{is}\right](I-A_i) \\ & + T^{-1}\sum_{t=2}^T Eu_{it}u_{it}' \end{aligned}$$

By Lemma 1, the maximum order of $\sum_{t=2}^{T}\left(\frac{1}{(t-1)^2}\right)E\left[\sum_{s=1}^{t-1}x_{is}\right]^2$ is $O\left(\sum_{t=2}^{T}\frac{1}{(t-1)}\right)$. Also, since $\sum_{t=1}^{T}\left(\frac{1}{t-1}\right)=\ln T+R_n$ where R_n is a residual, we have

$$T^{-1}\sum_{t=1}^{T}Ee_{it}e_{it}^{\prime}=O\left(T^{-1}\ln T\right)+T^{-1}\sum_{t=1}^{T}Eu_{it}u_{it}^{\prime}$$

It follows that e_{it} and u_{it} have the same asymptotic variance. ■

The $N-$asymptotic bias, variance and mean squared error formulae for RLSDV and LSDV are given in,

Claim 1 *($N-$Asymptotic Properties of RLSDV). Let the observations be generated by (10), and*

$$\begin{aligned}
A &= \frac{1}{T}\left[\frac{1}{1-\rho}\right]\left[T-\left(\frac{1-\rho^T}{1-\rho}\right)\right],\\
B &= \left[\frac{T-1}{1-\rho^2}\right]\left\{1-\left(\frac{1}{T-1}\right)\left(\frac{2\rho}{1-\rho}\right)\left[1-\left(\frac{1}{T}\right)\left(\frac{1-\rho^T}{1-\rho}\right)\right]\right\},\\
C &= \left(\frac{1}{1-\rho^2}\right)\sum_{t=2}^{T}\left(\frac{\rho}{t-1}\right)\left[1+\rho^t-\left(\frac{2}{t-1}\right)\left(\frac{1-\rho^{t-1}}{(1-\rho)}\right)\right],\\
D &= \left(\frac{1}{1-\rho^2}\right)\sum_{t=2}^{T}\left[1-\left(\frac{1}{t-1}\right)\left(1-\left(\frac{2\rho^{t-1}}{1-\rho}\right)+\left(\frac{2\rho}{t-1}\right)\left(\frac{1-\rho^{t-1}}{(1-\rho)^2}\right)\right)\right],\\
E &= \left(\frac{1}{T-1}\right)\sum_{t=2}^{T}\left(\frac{1}{t-1}\right)\left(1-\left(\frac{2\rho}{t-1}\right)\left(\frac{1-\rho^{t-1}}{1-\rho^2}\right)\right).
\end{aligned}$$

Then the $N-$asymptotic bias of the RLSDV estimator is

$$\underset{N\to\infty}{plim}\left(\rho_{RLSDV}-\rho\right)=\left(\frac{C}{D}\right)>0$$

Proof. The data generating process has the observationally equivalent latent model structure,

$$\begin{aligned}
q_{it} &= c_i+x_{it},\\
x_{it} &= \rho x_{it-1}+u_{it}.
\end{aligned}$$

For this latent structure, RPLS is defined as

$$\rho_{\mathrm{RLSDV}}=\rho+\left(\frac{1}{D_{NT}}\right)\{C_{1,NT}+C_{2,NT}\}$$

where

$$\begin{aligned}
C_{1,NT} &= \sum_{i=1}^{N}\sum_{t=2}^{T}\left(x_{it-1}-\frac{1}{t-1}\sum_{s=1}^{t-1}x_{is}\right)u_{it}, \\
C_{2,NT} &= -(1-\rho)\sum_{i=1}^{N}\sum_{t=2}^{T}\left(x_{it-1}-\frac{1}{t-1}\sum_{s=1}^{t-1}x_{is}\right)\left(\frac{1}{t-1}\sum_{s=1}^{t-1}x_{is}\right), \\
D_{NT} &= \sum_{i=1}^{N}\sum_{t=2}^{T}\left(x_{it-1}-\frac{1}{t-1}\sum_{s=1}^{t-1}x_{is}\right)^2 .
\end{aligned}$$

To evaluate the probability limit, $\operatorname*{plim}_{N\to\infty}(\rho_{\text{RLSDV}}-\rho)=\frac{\operatorname{plim}\frac{1}{N}\left[C_{1,NT}+C_{2,NT}\right]}{\operatorname{plim}\frac{1}{N}D_{NT}}$, we have

$$\begin{aligned}
\operatorname*{plim}_{N\to\infty}\left(\frac{1}{N}\right)D_{NT} &= E\sum_{t=2}^{T}\left(x_{it-1}-\frac{1}{t-1}\sum_{s=1}^{t-1}x_{is}\right)^2 \\
&= E\sum_{t=2}^{T}\left[x_{it-1}^2+\left(\frac{1}{t-1}\sum_{s=1}^{t-1}x_{is}\right)^2-2\frac{1}{t-1}x_{it-1}\sum_{s=1}^{t-1}x_{is}\right] \\
&= \sum_{t=2}^{T}\sigma_x^2+\sum_{t=2}^{T}\left(\frac{1}{t-1}\right)^2\left\{\sigma_x^2(t-1)\left[\frac{1+\rho}{1-\rho}-\frac{2\rho}{t-1}\frac{1-\rho^{t-1}}{(1-\rho)^2}\right]\right\} \\
&\quad -\left(\frac{2}{t-1}\right)\left(\frac{1-\rho^{t-1}}{1-\rho}\right)\sigma_x^2 \\
&= \sum_{t=2}^{T}\sigma_x^2\left(\begin{array}{c} 1+\left[\frac{1}{t-1}\right]\left[\frac{1+\rho}{1-\rho}\right]-\left[\frac{2}{t-1}\right]\left[\frac{1-\rho^{t-1}}{1-\rho}\right] \\ -\left[\frac{1}{t-1}\right]\left[\frac{2\rho}{t-1}\right]\left[\frac{1-\rho^{t-1}}{(1-\rho)^2}\right] \end{array}\right) \\
&= \sum_{t=2}^{T}\sigma_x^2\left[1-\left[\frac{1}{t-1}\right]\left(1-\left[\frac{2\rho^{t-1}}{1-\rho}\right]+\left[\frac{2\rho}{t-1}\right]\left[\frac{1-\rho^{t-1}}{(1-\rho)^2}\right]\right)\right]=D
\end{aligned}$$

where $\sigma_x^2=\operatorname*{plim}_{N\to\infty}\frac{1}{N}\sum\sigma_{i,x}^2$. Note that $\operatorname*{plim}_{N\to\infty}\frac{1}{N}C_{1,NT}=0$, whereas

$$
\begin{aligned}
\underset{N\to\infty}{\text{plim}}\frac{1}{N}C_{2,NT} &= -E\left(1-\rho\right)\sum_{t=2}^{T}\left[\left(x_{it-1}-\frac{1}{t-1}\sum_{s=1}^{t-1}x_{is}\right)\frac{1}{t-1}\sum_{s=1}^{t-1}x_{is}\right] \\
&= -E\left(1-\rho\right)\sum_{t=2}^{T}\left[x_{it-1}\frac{1}{t-1}\sum_{s=1}^{t-1}x_{is}-E\left(\frac{1}{t-1}\sum_{s=1}^{t-1}x_{is}\right)^{2}\right] \\
&= -\sum_{t=2}^{T}\left[\frac{1-\rho^{t-1}}{t-1}\right]\sigma_x^2+\sum_{t=2}^{T}\frac{\sigma_x^2}{t-1}\left(1+\rho-\frac{2\rho}{t-1}\frac{1-\rho^{t-1}}{1-\rho}\right) \\
&= \sigma_x^2\sum_{t=2}^{T}\frac{\rho}{t-1}\left(1+\rho^{t}-\frac{2}{t-1}\frac{1-\rho^{t-1}}{(1-\rho)}\right)=C.
\end{aligned}
$$

Finally we have the $N-$asymptotic bias of RLSDV,

$$
\underset{N\to\infty}{\text{plim}}\left(\rho_{\text{RLSDV}}-\rho\right)=\frac{\sum_{t=2}^{T}\left(\frac{\rho}{t-1}\right)\left(1+\rho^{t}-\left(\frac{2}{t-1}\right)\left(\frac{1-\rho^{t-1}}{(1-\rho)}\right)\right)}{(T-1)-\sum_{t=2}^{T}\left(\frac{1}{t-1}\right)\left(1-\left(\frac{2\rho^{t-1}}{1-\rho}\right)+\left(\frac{2\rho}{t-1}\right)\left(\frac{1-\rho^{t-1}}{(1-\rho)^2}\right)\right)}=\left(\frac{C}{D}\right).
$$

■

$N-$Asymptotic bias in panel AR(p) case

In section 2.3, we claimed that the RPLS bias may be more severe than the LSDV bias for the $\phi_j, j=2,\ldots p$ coefficients. The biases are

$$
\begin{aligned}
\underset{N\to\infty}{\text{plim}}\left(\underline{\phi}_{\text{LSDV}}-\underline{\phi}\right) &= \underset{N\to\infty}{\text{plim}}\left(\rho_{\text{LSDV}}-\rho\right)\left(\frac{\sum^{N}\sum^{T}z_{it}\tilde{y}_{it-1}}{\sum^{N}\sum^{T}\tilde{z}_{it}^{2}}\right)+\underset{N\to\infty}{\text{plim}}\left(\frac{\sum^{N}\sum^{T}\tilde{z}_{it}\tilde{u}_{it}}{\sum^{N}\sum^{T}\tilde{z}_{it}^{2}}\right) \\
\underset{N\to\infty}{\text{plim}}\left(\underline{\phi}_{\text{RLSDV}}-\underline{\phi}\right) &= \underset{N\to\infty}{\text{plim}}\left(\rho_{\text{RLSDV}}-\rho\right)\left(\frac{\sum^{N}\sum^{T}z_{it}\left(y_{it-1}-\mu_{it}\right)}{\sum^{N}\sum^{T}z_{it}^{2}}\right) \\
&\quad +\underset{N\to\infty}{\text{plim}}\left(\frac{\sum^{N}\sum^{T}z_{it}e_{it}}{\sum^{N}\sum^{T}z_{it}^{2}}\right)
\end{aligned}
$$

Residual bias in second stage ϕ estimates for panel AR(p)

In section 2.3, it was claimed that the residual bias of the estimator $\hat{\underline{\phi}}^{\dagger}$ is inconsequential. To fix the idea, we take AR(2) as an example and rewrite (13) as

$$
y_{it}-\rho_{\text{RLSDV}}y_{it-1}=\alpha_i+\phi\left(y_{it-1}-y_{it-2}\right)+u_{it}^{\dagger} \tag{17}
$$

where $u_{it}^{\dagger} = u_{it} - (\rho_{\text{RLSDV}} - \rho)\, y_{it-1}$. Using the fact that $(y_{it-1} - \mu) - (y_{it-2} - \mu) = (y_{it-1} - y_{it-2})$ where μ is unconditional mean of y_{it}, the bias of the pooled estimator $\underline{\hat{\phi}}^{\dagger}$ can be written as

$$\underline{\hat{\phi}}^{\dagger} - \underline{\phi} = (\rho - \rho_{\text{RLSDV}}) \frac{\sum^N \sum^T (y_{it-1} - y_{it-2})(y_{it-1} - y_{i\cdot-1})}{\sum^N \sum^T (y_{it-1} - y_{it-2})^2} + \frac{\sum^N \sum^T (y_{it-1} - y_{it-2})(u_{it} - u_{i\cdot})}{\sum^N \sum^T (y_{it-1} - y_{it-2})^2}$$

As $N \to \infty$, the asymptotic bias is given by

$$\operatorname*{plim}_{N\to\infty} \left(\underline{\hat{\phi}}^{\dagger} - \underline{\phi}\right) = \operatorname*{plim}_{N\to\infty} (\rho - \rho_{\text{RLSDV}})\, O(1) + O\left(T^{-1}\right) \operatorname*{plim}_{N\to\infty} \frac{1}{N} \sum^N \sum^T (y_{it-1} - y_{it-2})(u_{it} - u_{i\cdot})$$

However note that

$$\left(1 - \rho_1 L - \rho_2 L^2\right) y_{it} = (1 - \lambda_1 L)(1 - \lambda_2 L)\, y_{it} = u_{it}$$

or

$$y_{it} = \frac{c_1}{1 - \lambda_1 L} u_{it} + \frac{c_2}{1 - \lambda_2 L} u_{it}$$

Since

$$\begin{aligned} & E\frac{1}{T}\left(\sum_{j=0}^{\infty} c_s \lambda_s^j u_{it-j}\right)\left(\sum_{t=1}^{T} u_{it}\right) - E\frac{1}{T}\left(\sum_{j=0}^{\infty} c_s \lambda_s^j u_{it-j-1}\right)\left(\sum_{t=1}^{T} u_{it}\right) \\ = \; & c_s (1 - \lambda_1) - c_s \left(\frac{T-1}{T}\right)(1 - \lambda_1) + O\left(T^{-2}\right) = O\left(T^{-1}\right) \quad \text{for } s = 1, 2. \end{aligned}$$

where

$$c_1 = \lambda_1 (\lambda_1 - \lambda_2) \text{ and } c_2 = -\lambda_2 / (\lambda_1 - \lambda_2)\,.$$

Hence we have

$$\begin{aligned} & \operatorname*{plim}_{N\to\infty} \frac{1}{N} \sum^N \sum^T (y_{it-1} - y_{it-2})(u_{it} - u_{i\cdot}) \\ = \; & -\text{plim} \frac{1}{N} \sum^N \sum^T (y_{it-1} - y_{it-2})\, u_{i\cdot} = O\left(T^{-1}\right) \end{aligned}$$

That is,

$$\operatorname*{plim}_{N\to\infty} \left(\underline{\hat{\phi}}^{\dagger} - \underline{\phi}\right) = \operatorname*{plim}_{N\to\infty} \left(\rho - \hat{\rho}^{\text{RPLS}}\right) O(1) + O\left(T^{-2}\right)$$

The above logic goes through for general AR(p), hence we can say that

$$\begin{aligned} \operatorname*{plim}_{N\to\infty} \left(\underline{\hat{\phi}}^{\dagger} - \underline{\phi}\right) &= \operatorname*{plim}_{N\to\infty} (\rho - \rho_{\text{RLSDV}}) \left(\frac{\sum^N \sum^T z_{it} y_{it-1}}{\sum^N \sum^T z_{it}^2}\right) + \operatorname*{plim}_{N\to\infty} \left(\frac{\sum^N \sum^T z_{it} \tilde{u}_{it}}{\sum^N \sum^T z_{it}^2}\right) \\ &= \operatorname*{plim}_{N\to\infty} (\rho - \rho_{\text{RLSDV}})\, O(1) + O\left(T^{-2}\right) \end{aligned}$$

it follows that the residual bias in $\underline{\hat{\phi}}^\dagger$ is inconsequential.