NBER WORKING PAPER SERIES

UNBIASED ESTIMATION OF THE HALF-LIFE TO PPP CONVERGENCE IN PANEL DATA

Chi-Young Choi
Nelson C. Mark
Donggyu Sul

Working Paper 10614
http://www.nber.org/papers/w10614

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
June 2004

Unbiased Estimation of the Half-Life to PPP Convergence in Panel Data
Chi-Young Choi, Nelson C. Mark, and Donggyu Sul
NBER Working Paper No. 10614
June 2004
JEL No.C32, F31, F47

ABSTRACT

Three potential sources of bias present complications for estimating the half-life of purchasing power parity deviations from panel data. They are the bias associated with inapproriate aggregation across heterogeneous coefficients, time aggregation of commodity prices, and downward bias in estimation of dynamic lag coefficients. Each of these biases have been addressed individually in the literature. In this paper, we address all three biases in arriving at our estimates. Analyzing an annual panel data set of real exchange rates for 21 OECD countries from 1948 to 2002, our point estimate of the half-life is 5.5 years.

Chi-Young Choi
Department of Economics
University of New Hampshire
cy.choi@unh.edu

Nelson Mark
Department of Economics and Econometrics
University of Notre Dame
Notre Dame, IN 46556
and NBER
nmark@nd.edu

Donggyu Sul
Department of Economics
University of Auckland
dsul@auckland.ac.nz

1 Introduction

The motivation for using panel data to estimate the convergence half-life of purchasing power parity (PPP) deviations is straightforward. Increasing the number of data points by combining the cross-section with the time-series should give more precise estimates.[1] Obtaining accurate measurements of the convergence speed to PPP is important because of its role in guiding theoretical work on the role of nominal rigidities and the relative importance of nominal and real shocks in macro models. Accuracy in estimation is especially important due to the nonlinearity of the half-life formula as small differences in the estimated value of dynamic lag coefficients of the real exchange rate can lead to markedly different predictions of the half life.

In practice, panel data estimation of the half-life to convergence has been anything but straightforward. Popular estimators are potentially subject to three sources of bias. Each of these biases have been discussed in the literature and addressed by researchers on an individual basis. In this short paper, we address and control for all three potential sources of bias to arrive at our final estimate of the half life. Using an annual sample of 21 OECD country's CPI-based real exchange rates from 1973 to 2002, when we jointly control for multiple sources of bias, we estimate the half-life to be 5.5 years (95 percent confidence interval ranges from 4.3 to 7.3 years). This approximately brings us back to point estimates obtained by the uncorrected least-squares dummy variable method. It is also within Rogoff's (1996) consensus estimate of the half life so the PPP puzzle remains in panel data.[2]

One potential source of bias is inappropriate pooling across cross sectional units. If the real exchange rates of different countries exhibit heterogenous rates of convergence to PPP then the data should not be pooled because the panel data estimator of a common autoregressive coefficient can be biased upwards. Imbs et. al. (2004) study how sectoral heterogeneity in convergence rates to the law of one price can result in upward bias of the estimated half-life but Chen and Engel (2004) find that sectoral heterogeneity is not a quantitatively important source of bias. We do not address sectoral heterogeneity in this paper but we do examine the potential importance of country-level heterogeneity.

A second source of bias is that price indices used to form real exchange rates are not constructed from point-in-time sampled commodity prices. Instead, source agencies report period averages of commodity and service prices. The consequences of this time

[1]Frankel and Rose (1996) was one of the first PPP studies to use panel data. Panel data analysis has been useful in forming a concensus that PPP holds in the long run. While univariate tests on post-1973 data generally cannot reject a unit root in the real exchange rate, panel unit root tests provide consistent rejections of the unit root hypothesis. See Chiu (2002), Choi (2002), Fleissig and Strauss (2000), Flores et. al. (1999), Lothian (1997), Papell and Theodoridis (1998), and Papell (2004). The alternative is to obtain long historical time series, as in Lothian and Taylor (1996) but because those data span a variety of regimes they pose their own set of complications.

[2]The PPP puzzle is that real exchange rates exhibit both very long half-lives (three to five years) and high short-term volatility.

aggregation of the data was first discussed by Working (1960). Taylor (2001) extended the analysis to the study of PPP by showing that time-aggregation biases results in an upward bias in the estimated half-life.

The third source of bias that we consider is the attendant downward small sample bias that results when a constant is included in the dynamic regression. This bias was discussed in the univariate context by Marriott and Pope (1954) and Kendall (1954), and in the dynamic panel context by Nickell (1981). The source of this downward bias can be seen by noting that using least squares to estimate an autoregression with a constant is equivalent to running the regression with no constant on observations that are deviations from the sample mean. The problem then, is that for any observation t, the regression error is correlated with current and future values of the real exchange rate which are embedded in the sample mean which in turn is a component of the independent variable. It is this induced correlation between the regression error and the sample mean that gives rise to the downward bias. Allowing for fixed effects, the half life based on the least-squares dummy variable (LSDV) estimator of ρ will be biased down and will give estimates of half lives that are too short.[3]

The remainder of the paper is organized as follows. The next section discusses our measurement of the half-life. Section 3 discusses each of the potential biases. Section 4 outlines our bias-adjustment strategies and presents the empirical results. Section 5 concludes. An appendix contains derivations for many of the results presented in the text.

2 Half life measurement

Let the real exchange rate for country i, $(i = 1, ..., N)$ evolve according to a first-order autoregression (AR(1)), $q_{it} = \alpha_i + \rho q_{it-1} + e_{it}$, where e_{it} is serially uncorrelated. The half-life $H(\rho)$, commonly employed as a measure of the speed at which convergence to PPP occurs, is the time required for a unit shock to PPP to dissipate by one half. In the AR(1) case, it is t^* such that $E(q_{t^*}) = e_1/2 = 1/2$, which takes the convenient form

$$t^* = H\left(\rho\right) = \frac{\ln(0.5)}{\ln\left(\rho\right)} \tag{1}$$

Due to the nonlinear nature of $H(\rho)$, small variations in ρ lead to disproportionately large variations in the half life, especially for ρ in the region near unity.[4] Thus if the estimator of ρ is biased, failure to provide appropriate adjustments can produce substantially misleading estimates of the half life.

For more complicated dynamic models that include additional lags or moving average

[3] The LSDV method is pooled OLS with fixed effects. See Hsiao (2003).

[4] e.g., $H(0.93) = 9.56, H(0.95) = 13.5, H(0.97) = 22.8$.

error terms, eq.(1) would only approximate the true half life. For these models, the exact half life can be computed by impulse response analysis. A knotty problem associated with general ARMA specifications is that the impulse response may not be monotonic so that there may be multiple half lives. We are able to avoid these complications in our empirical analysis by employing annual data for which the AR(1) specification is appropriate.

3 Three possible sources of half-life bias

In this section, we review the three potential sources of bias discussed in the literature. Section 4 discusses our strategy for accounting for these biases.

Cross-sectional aggregation bias. Imbs et. al. (2004) study how heterogeneity in the speed of convergence towards the law of one price across the different commodities that comprise the general price level may result in an upward bias of the estimated half-life of PPP deviations. Chen and Engel (2004), on the other hand, find that sector heterogeneity is not a quantitatively important source of bias. We will not address the issue of sectoral heterogeneity but we do consider the possibility that there is country specific heterogeneity in convergence rates.

To see how cross-sectional heterogeneity can bias the panel estimator, suppose that the real exchange rate for country i follows[5]

$$q_{it} = \rho_i q_{it-1} + e_{it}. \tag{2}$$

If the heterogeneity in the autoregressive coefficient across countries is specified as

$$\rho_i = \rho + v_i \tag{3}$$

where $E(v_i) = 0$, then substituting (3) into (2) gives

$$q_{it} = \rho q_{it-1} + (e_{it} + v_i q_{it-1}). \tag{4}$$

The potential bias arises because the second piece of the composite error term $v_i q_{it-1}$, is correlated with the regressor $q_{it-1.}$ Looking in more detail at the pooled OLS estimator,

$$\widehat{\rho}_{OLS} = \rho + \underbrace{\frac{\sum_{i=1}^{N}\sum_{t=1}^{T} q_{it-1}e_{it}}{\sum_{i=1}^{N}\sum_{t=1}^{T} q_{it-1}^2}}_{A(N,T)} + \underbrace{\frac{\sum_{i=1}^{N} v_i \left(\sum_{t=1}^{T} q_{it-1}^2\right)}{\sum_{i=1}^{N}\sum_{t=1}^{T} q_{it-1}^2}}_{B(N,T)} \tag{5}$$

[5] We disregard the constant here so as to isolate the bias arising from cross-sectional aggregation.

$A(N,T)$ is standard. The piece introduced by aggregating across heterogeneous cross-sectional coefficients is $B(N,T)$. If each of the country real exchange rates are covariance stationary and the distribution of v_i is symmetric, then this is unlikely to be a quantitatively important source of bias because the terms $v_i\left(\sum_{t=1}^{T} q_{it-1}^2\right)$ will average out to zero. What is potentially a more serious situation is if the observations are drawn from a mixed panel where a fraction π of the real exchange rates are stationary and a fraction $1-\pi$ are unit root nonstationary. In this case, the OLS estimator can be shown to be

$$\widehat{\rho}_{OLS} = \frac{\rho\pi\left(\sum_{i=1}^{N}\frac{1}{\left(1-\rho_i^2\right)}\right) + (1-\pi)\left(\frac{T+1}{2}\right)}{\pi\left(\sum_{i=1}^{N}\frac{1}{\left(1-\rho_i^2\right)}\right) + (1-\pi)\left(\frac{T+1}{2}\right)} \geq \rho \tag{6}$$

which evidently is biased upwards.[6] If there is heterogeneity in the data, pooling is inappropriate and an alternative estimation strategy should be employed.

Nickell bias. Nickell (1981) studied the properties of the LSDV estimator for the dynamic panel regression model when the observations are cross-sectionally independent. His analysis showed that pooling results in more efficient estimates of ρ than OLS but does not eliminate the downward bias found in univariate estimation. The bias in the LSDV estimator does not go away even asymptotically (when $N\to\infty$). We refer to this as $N-$asymptotic bias.

For the LSDV estimator in the panel AR(1) model with fixed effects, Nickell shows

$$\begin{aligned} \operatorname*{plim}_{N\to\infty}\hat{\rho}_{LSDV} &\equiv m(\rho) \\ &= \rho - \left(\frac{1+\rho}{T-1}\right)\left[1-\left(\frac{1}{T}\right)\left(\frac{1-\rho^T}{1-\rho}\right)\right] \\ &\quad \times\left\{1-\left(\frac{1}{T-1}\right)\left(\frac{2\rho}{1-\rho}\right)\left[1-\left(\frac{1}{T}\right)\left(\frac{1-\rho^T}{1-\rho}\right)\right]\right\}^{-1} \end{aligned} \tag{7}$$

which is biased downwards.

Time aggregation bias. Time aggregation bias was first analyzed by Working (1960) and subsequently studied by many authors.[7] Working showed that if the true underlying process followed a driftless random walk, that time-averaging of this process induces a moving average error into the reported (time-averaged) first differences. The analyst who

[6] Derivation in Appendix section 1.

[7] Tiao (1972) and Brewer (1973) also develop econometric implications of time aggregation. Rossana and Seater (1995), Marcellino (1999), and Breitung and Swanson (1999) study the effects of time aggregation on exogeneity tests and forecasting.

estimates the correlation of first-differenced time-averaged observations will mistakenly conclude that they are serially correlated when in fact the autocorrelation is zero. Taylor (2001) extends this to the case where the true process follows a stationary AR(1). In the PPP problem, an upward bias is induced in estimation of ρ because source statistical agencies report price indices that are formed as averages of goods and services prices over a particular interval and are not point-in-time sampled prices. He reports that this is standard practice around the world and argues that the 3-5 year consensus half life overstates the truth because those studies did not correct for time-aggregation bias.

With time aggregated observations, the data are reported at time intervals indexed by t but within each data reporting interval there are M subintervals at which the underlying price process is observable. Thus if the data are reported annually, there are $M = 260$ business days and the annual observations are reported as period averages at the annual time intervals $j = M, 2M, ..., TM$. Assuming that the dynamics of the underlying point-in-time daily real exchange rate process evolves according to the AR(1) process $q_{ij} = a_i + \phi q_{ij-1} + e_{ij}$, with autocorrelation coefficient ϕ, the dynamics of the point sampled process at annual intervals is $q_{ij+M} = \alpha_i + \phi^M q_{ij} + e_{ij+M}$ with autocorrelation coefficient $\phi^M < \phi$ for $0 < \phi < 1$ and the 'true' half life in years is $H(\phi) = \ln(0.5)/\ln(\phi^M)$. However, when the available observations are the average of prices over $M = 260$ working days, the data being analyzed are $\frac{1}{M}\sum_{j=1}^{M} q_{i,Mt-j}$. Taylor shows that the implied autocorrelation coefficient from fitting the time-averaged annual real exchange rate to an AR(1) is

$$\rho \equiv G(\phi, M) = \frac{\phi(1-\phi^M)^2}{M(1-\phi^2) - 2\phi(1-\phi^M)} > \phi^M \tag{8}$$

which leads to an overstatement of the half-life.

Since point sampled nominal exchange rates are available, one might be tempted to combine them with the time-averaged price indices to mitigate time aggregation bias embedded in the real exchange rate. However, nuisance parameter dependencies make it impossible to determine the bias in the combined point and time-averaged data. A discussion of this issue is provided in the appendix

4 Bias-adjusted half-life estimation

The data. We use annual real exchange rates of 21 OECD countries which are constructed by combining annual nominal exchange rates and annual consumer price indices from 1948 to 1998 which results in 51 observations. Both nominal exchange rates (IFS line code RF) and CPIs (IFS line code 64) are annual average observations. They were retrieved from the International Monetary Fund's International Financial Statistics (IFS) for 21 industrial countries: Australia, Austria, Belgium, Canada, Denmark, Finland, France, Germany, Greece, Ireland, Italy, Japan, the Netherlands, New Zealand, Norway, Portugal, Spain, Sweden, Switzerland, the United Kingdom, and the United States. Each country is alternatively considered as the numeraire country.[8]

In preliminary data analysis, we employed the Phillips and Sul (2003) panel unit root test which finds that the real exchange rates defined by the alternative numeraires are stationary. We do not devote space for detail reporting of these results since they simply confirm the findings of earlier research.

Cross-sectional heterogeneity. We investigate whether pooling is appropriate in our data set. In order for the test of the homogeneity restrictions to have the correct size, the test must be done using an estimator that controls for Nickell bias. For this purpose, we estimate systems associated with each of the numeraire countries by recursive mean adjusted seemingly unrelated regression and conduct tests for homogeneity of ρ.[9] The results of the homogeneity test are reported in Table 1. Homogeneity is rejected at the 5 percent level only when Belgium, France, and Greece are the numeraire countries. Because the evidence against homogeneity in this data set is fairly weak, we proceed by assuming that pooling is generally appropriate and that cross-country heterogeneity in the autoregressive coefficient does not constitute a significant source of bias.

Table 1: Homogeneity Test (Real Exchange Rates in 21 OECD Countries 1948-1998)

Numeraire Country	Wald Statistic	P-value	Numeraire Country	Wald Statistic	P-value
Australia	14.02	0.78	Japan	23.60	0.21
Austria	20.15	0.39	Netherlands	16.56	0.62
Belgium	33.82	0.02*	New Zealand	33.80	0.02
Canada	24.46	0.18	Norway	11.40	0.91
Denmark	14.98	0.72	Portugal	18.83	0.47
Finland	22.62	0.25	Spain	8.21	0.98
France	43.96	0.00*	Sweden	8.94	0.97
Germany	15.83	0.67	Switzerland	17.81	0.53
Greece	33.82	0.02*	U.K.	26.76	0.11
Ireland	13.73	0.80	U.S.	26.74	0.11
Italy	11.22	0.92			

[8] Papell and Theodoridis (2001) find that the choice of numeriare matters in panel unit root tests of PPP.

[9] See Choi, Mark and Sul (2004) for a description of this estimator and its properties.

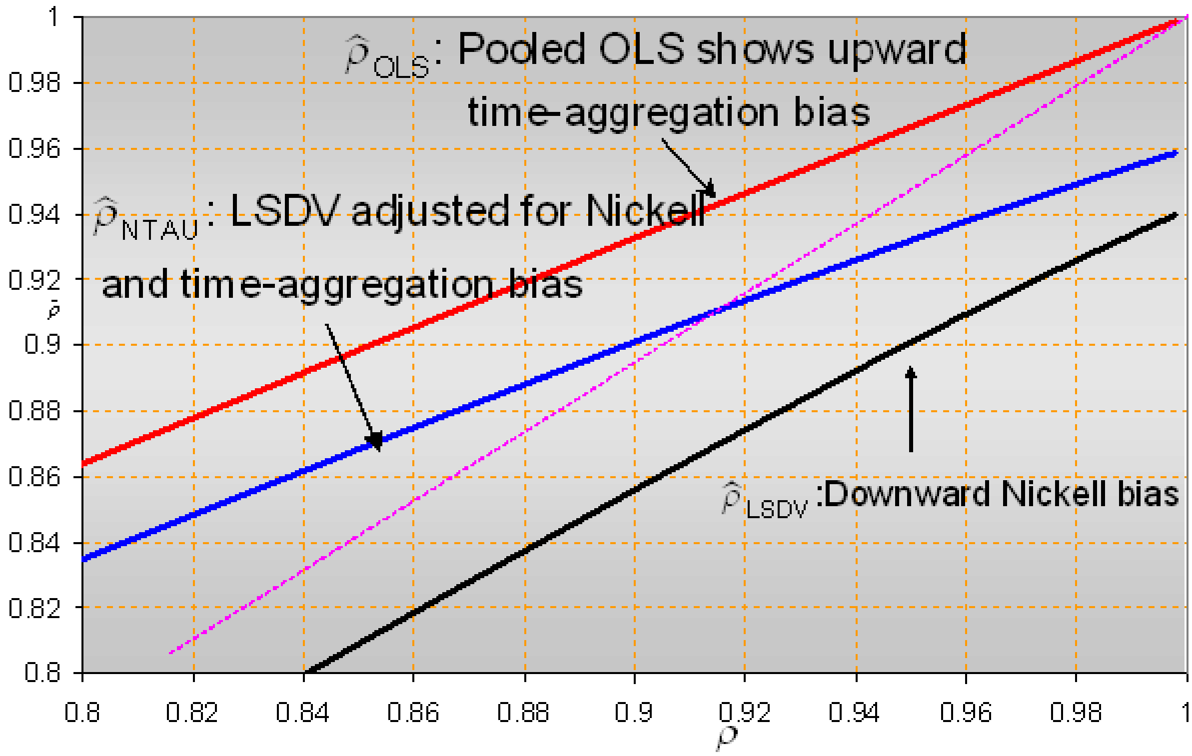

Figure 1: $N-$asymptotic biases of pooled estimators with pure temporal aggregated data (T=51).

Combined Nickell and time-aggregation bias adjustments We turn now to a joint treatment of Nickell bias and time-aggregation bias. If we had point-sampled data so that time aggregation bias were not an issue, the adjustment for Nickell bias can be done directly with panel mean unbiased estimation. This would proceed by estimating ρ by LSDV and then adjusting for bias using the inverse function of the $N-$asymptotic bias formula to obtain the mean unbiased estimator $\hat{\rho}_{MUE}$ as the inverse of the mean function, $\hat{\rho}_{MUE} = m^{-1}(\hat{\rho}_{LSDV})$ where $\widehat{\rho}_{LSDV}$ is the LSDV estimator and the function $m(\cdot)$ is given in (**??**).[10] However, when the data are time aggregated, a further adjustment in the mean function is necessary to do panel mean unbiased estimation as there is now an interaction between Nickell's fixed effect bias and the time aggregation bias.

To better understand the relation between the two biases, we are able to analytically

[10]Murray and Papell (2002)employed median unbiased estimation whose performance is very similar to panel mean unbiased estimation.

characterize the LSDV bias with time aggregated data under cross section independence although in estimation we will relax the independence assumption. The pure Nickell bias and the time-aggregation bias go in opposite directions and a decomposition of the opposing bias factors is shown in Figure 1. In the figure, the true value of ρ is plotted on the horizontal axis and the LSDV probability is plotted on the vertical axis.[11] The top line shows the effect of time-aggregation in panel data. It is the probability limit of the pooled OLS estimator on time aggregated data with no regression constant. In this case, a pooled OLS point estimate of 0.9 (implied half life of 6.6 years) implies that the time-aggregated bias corrected value of ρ is approximately 0.85 (implied half life of 4.3 years). As $\rho \to 1$, the upward time aggregation bias vanishes. The bottom line shows the effect of pure Nickell bias which is the LSDV probability limit from (**??**). For this case, an LSDV point estimate of 0.9 implies that the mean-unbiased value of ρ is approximately 0.95 (implied half life of 13.5 years). The center line shows the effects of the combined biases. In the neighborhood of $\rho = 0.9$, the two pieces largely offset each other. When the true value of ρ lies below (above) 0.9, however there is an upward (downward) combined bias.

Denote the formula that generates the center line by $B(\rho, T)$ (shown in the appendix). A strategy that simultaneously corrects for Nickell and time aggregation bias is to estimate ρ by LSDV and invert the function,[12]

$$\widehat{\rho}_{\text{NTAU}} = B^{-1}\left(\rho_{LSDV}, T\right). \tag{9}$$

To this proposed correction, we make one additional adjustment. Because LSDV does not exploit the cross-sectional covariance structure of the observations in estimation, an efficiency improvement can be achieved by using a panel GLS estimator with fixed effects. When the cross-sectional dependence has a single factor structure, the $N-$asymptotic bias of the fixed effects GLS estimator is independent of both the factor loadings and the unobserved factor (Phillips and Sul (2003)). This independence allows us to apply the mean adjustment in (9) with the panel GLS estimator in place of LSDV. Call it $\widehat{\rho}_{\text{GNTAU}}$. A brief description of the estimator is given in the appendix.

Table 2 reports our panel estimates of ρ and implied median and 95 percentiles for the half lives of all 21 panels defined by alternative numeraire. The unadjusted LSDV estimate with the US as numeraire is $\widehat{\rho}_{\text{LSDV}} = 0.912$ implies a half life of 7.5 years. The median half life across all 21 panels is 5.6 years. Applying the Nickell mean adjustment (ignoring time aggregation) gives the mean-unbiased estimate $\widehat{\rho}_{\text{MUE}} = 0.96$ when the US is the numeraire country. This gives an implied half life of 17 years. The median half life across all numeraire country panels is 22 years. Finally, when we jointly adjust

[11] The probability limits are for $N \to \infty$ but for fixed $T = 51$ which corresponds to the number of time series observations in our data set.

[12] The Nickell and Time Aggegation Unbiased estimator.

for Nickell bias and time aggregation bias, estimating the autocorrelation coefficient by GLS and applying the mean adjustment, we obtain an estimate of the autocorrelation coefficient with the US as numeraire of $\widehat{\rho}_{\text{GNTAU}} = 0.87$ which is slightly below the LSDV estimate. The implied half life is 4.8 years (95 percent confidence interval ranges from 3.7 to 6.6 years). The median across all alternative panels is 5.5 years which approximately returns us to the LSDV estimates.

5 Conclusion

PPP research, desperate for larger sample sizes to improve precision and confidence in empirical estimates, has turned to the analysis of panel data. However, half-life estimation from panel data, is subject to two major sources of bias. The first source is the downward bias of the LSDV estimator. The second main source of bias arises from the use of time aggregated data. These two opposing biases roughly cancel out. When we simultaneously control for these two opposing biases, the resulting half-life estimates bring us approximately back to the LSDV estimates.

Table 2: Panel Half Life Estimation

Numeraire	$\widehat{\rho}_{\text{LSDV}}$	$H_{0.025}$	$H_{0.5}$	$H_{0.975}$	$\widehat{\rho}_{\text{MUE}}$	$H_{0.5}$	$\widehat{\rho}_{\text{GNTAU}}$	$H_{0.025}$	$H_{0.5}$	$H_{0.975}$
Australia	0.890	4.6	5.9	8.1	0.977	29	0.884	4.3	5.6	7.8
Austria	0.866	3.9	4.8	6.3	0.970	23	0.883	4.4	5.6	7.3
Belgium	0.883	4.5	5.6	7.3	1.000	oo	0.941	7.8	11.4	21.2
Canada	0.948	9.1	13.1	23.1	0.957	16	0.828	3.0	3.7	4.7
Denmark	0.916	6.1	7.9	10.9	0.955	15	0.853	3.5	4.3	5.6
Finland	0.788	2.4	2.9	3.6	0.963	18	0.926	6.6	9.0	14.2
France	0.884	4.5	5.6	7.4	0.985	45	0.852	3.4	4.3	5.7
Germany	0.861	3.8	4.6	5.9	0.971	23	0.930	6.8	9.5	15.6
Greece	0.790	2.6	2.9	3.4	0.973	25	0.922	6.6	8.5	11.9
Ireland	0.872	4.1	5.1	6.6	0.969	22	0.906	5.2	7.1	10.9
Italy	0.881	4.4	5.5	7.2	0.961	17	0.845	3.3	4.1	5.4
Japan	0.962	12.3	17.7	31.2	0.947	13	0.811	2.7	3.3	4.2
Netherlands	0.912	5.9	7.5	10.3	0.962	18	0.868	3.9	4.9	6.5
New Zealand	0.862	3.8	4.7	6.0	0.971	23	0.882	4.3	5.5	7.7
Norway	0.905	5.5	6.9	9.1	0.970	23	0.877	4.3	5.3	6.8
Portugal	0.895	4.9	6.3	8.5	0.971	23	0.859	3.6	4.6	6.2
Spain	0.854	3.6	4.4	5.7	0.969	22	0.931	6.9	9.7	16.5
Sweden	0.887	4.6	5.8	7.8	0.973	26	0.919	5.9	8.2	13.1
Switzerland	0.920	6.4	8.3	11.6	0.966	20	0.869	3.9	4.9	6.6
U.K.	0.868	4.0	4.9	6.4	0.968	22	0.905	4.9	6.9	11.8
U.S.	0.912	5.8	7.5	10.5	0.960	17	0.865	3.7	4.8	6.6
Median	0.884	4.5	5.6	7.4	0.969	22	0.882	4.3	5.5	7.3

References

[1] Brewer, K.R.W., 1973. "Some Consequences of Temporal Aggregation and Systematic Sampling for ARMA and ARMAX Models," *Journal of Econometrics* 1, 133-154.

[2] Breitung, Jorg and Norman R. Swanson, 2002. "Temporal Aggregation and Spurious Instantaneous Causality in Multiple Time Series Models," *Journal of Time Series Analysis,* 23, 651-665.

[3] Chiu, Ru-Lin, 2002. "Testing the purchasing power parity in panel data," *International Review of Economics and Finance*, 11, 349–362.

[4] Choi, Chi-Young., 2002. "Searching for Evidence of Long-Run PPP from a Post Bretton Woods Panel: Separating the Wheat from the Chaff," *Journal of International Money and Finance, forthcoming.*

[5] Choi, Chi-Young, Nelson C. Mark and Donggyu Sul, 2004. "Bias Reduction by Recursive Mean Adjustment in Panel Data," *mimeo,* University of Auckland.

[6] Chen, Shiu-Sheng and Charles Engel, 2004. "Does 'Aggregation Bias' Explain the PPP Puzzle?" *mimeo,* University of Wisconsin.

[7] Fleissig, Adrian R. and Jack Strauss, 2000. "Panel Unit Root Tests of Purchasing Power Parity for Price Indices," *Journal of International Money and Finance* 19, 489–506.

[8] Flores, Renato, Philippe Jorion, Pierre-Yves Preumont, and Ariane Szafarz, 1999. "Multivariate unit root tests of the PPP hypothesis," *Journal of Empirical Finance*, 6, 335–353.

[9] 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.

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

[11] Imbs, J., Mumtaz, H., Ravn, M.O., Rey, H., 2004, "PPP Strikes Back: Aggregation and the Real Exchange Rate," *mimeo*, London Business School.

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

[13] Lothian, James R. 1997. "Multi-Country Evidence on the Behavior of Purchasing Power Parity under the Current Float," *Journal of International Money and Finance*, 16, 19–35.

[14] Lothian, James R. and Mark P. Taylor, 1996. "Real Exchange Rate Behavior: The Recent Float from the Perspective of the Past Two Centuries," *Journal of Political Economy*, 104, 488-509.

[15] Marcellino, Massimiliano, 1999. "Some Consequences of Temporal Aggregation in Empirical Analysis," *Journal of Business and Economic Statistics*, 17(1), 129–136.

[16] Marriott, F.H.C., Pope, J.A., 1954. "Bias in the Estimation of Autocorrelations," *Biometrika*, 41, 393-403

[17] Murray, Christian.J., Papell, David.H., 2002. "The Purchasing Power Parity Persistence Paradigm," *Journal of International Economics* 56(1), 1–19.

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

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

[20] Papell, David H., and Hristos Theodoridis, 1998. "Increasing Evidence of Purchasing Power Parity over the Current Float," *Journal of International Money and Finance*, 17, 41-50.

[21] Papell, David H., and Hristos Theodoridis, 2001. "The Choice of Numeraire Currency in Panel Tests of Purchasing Power Parity," *Journal of Money, Credit, and Banking,* 33, 790-803.

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

[23] Rogoff, Kenneth., 1996. "The Purchasing Power Parity Puzzle," *Journal of Economic Literature* 34(2), 647–668.

[24] Rossana, R.J. and Seater J.J. 1992. "Aggregation, Unit Roots and the Time-Series Structure of Manufacturing Real Wages," *International Economic Review*, 33(1), 159–179.

[25] Taylor, A.M., 2001. "Potential Pitfalls for the PPP Puzzle? Sampling and Specification Biases in Mean-Reversion Tests of the Law of One Price." *Econometrica* 69(2), 473–498.

[26] Tiao, G. C., 1972. "Asymptotic Behavior of Temporal Aggregates of Time Series,"" *Biometrika,* 59, 525-531.

[27] Working, Holbrook., 1960. "Note on the Correlation of First Differences of Averages in a Random Chain," *Econometrica* 28(4), 916-918.

Appendix

Derivation of equation (6)

$$
\begin{aligned}
\underset{N\to\infty}{\text{plim}}\hat{\rho} &= \underset{N\to\infty}{\text{plim}}\left(\frac{\frac{1}{N}\sum_{i=1}^{N}\sum_{t=1}^{T}q_{it}q_{it-1}}{\frac{1}{N}\sum_{i=1}^{N}\sum_{t=1}^{T}q_{it-1}^{2}}\right)\\
&= \underset{N\to\infty}{\text{plim}}\left[\frac{\frac{1}{N}\left(\sum_{i=1}^{N_1}\sum_{t=1}^{T}q_{it}^{s}q_{it-1}^{s}+\sum_{i=1}^{N_2}\sum_{t=1}^{T}q_{it}^{N}q_{it-1}^{N}\right)}{\frac{1}{N}\left(\sum_{i=1}^{N_1}\sum_{t=1}^{T}\left(q_{it-1}^{s}\right)^2+\sum_{i=1}^{N_2}\sum_{t=1}^{T}\left(q_{it-1}^{N}\right)^2\right)}\right]\\
&= \underset{N\to\infty}{\text{plim}}\left[\frac{\frac{N_1}{N}\frac{1}{N_1}\sum_{i=1}^{N_1}\sum_{t=1}^{T}q_{it}^{s}q_{it-1}^{s}+\frac{N_2}{N}\frac{1}{N_2}\sum_{i=1}^{N_2}\sum_{t=1}^{T}q_{it}^{N}q_{it-1}^{N}}{\frac{N_1}{N}\frac{1}{N_1}\sum_{i=1}^{N_1}\sum_{t=1}^{T}\left(q_{it-1}^{s}\right)^2+\frac{N_2}{N}\frac{1}{N_2}\sum_{i=1}^{N_2}\sum_{t=1}^{T}\left(q_{it-1}^{N}\right)^2}\right]\\
&= \frac{\pi\text{plim}_{N_1\to\infty}\left(\frac{1}{N_1}\sum_{i=1}^{N_1}\sum_{t=1}^{T}q_{it}^{s}q_{it-1}^{s}\right)+(1-\pi)\,\text{plim}_{N_2\to\infty}\left(\frac{1}{N_2}\sum_{i=1}^{N_2}\sum_{t=1}^{T}q_{it}^{N}q_{it-1}^{N}\right)}{\pi\text{plim}_{N_1\to\infty}\left(\frac{1}{N_1}\sum_{i=1}^{N_1}\sum_{t=1}^{T}\left(q_{it-1}^{s}\right)^2\right)+(1-\pi)\,\text{plim}_{N_2\to\infty}\left(\frac{1}{N_2}\sum_{i=1}^{N_2}\sum_{t=1}^{T}\left(q_{it-1}^{N}\right)^2\right)}
\end{aligned}
$$

Note that N_1 and $N_2 \to \infty$ as $N \to \infty$ since $\pi = N_1/N$ is a fixed constant and $\rho_i = \rho + \mu_i$ where $\mu_i - iidN\left(0, \sigma_\mu^2\right)$. Define $\lambda = \text{plim}_{N_1\to\infty}\frac{1}{N_1}\sum_{i=1}^{N_1}\frac{1}{1-\rho_i^2} < \infty$.

$$
\begin{aligned}
\underset{N_2\to\infty}{\text{plim}}\left(\frac{1}{N_2}\sum_{i=1}^{N_2}\sum_{t=1}^{T}q_{it}^{N}q_{it-1}^{N}\right) &= \sigma_e^2\sum_{t=1}^{T}t = \sigma_e^2\frac{T\left(T+1\right)}{2}\\
\underset{N_1\to\infty}{\text{plim}}\left(\frac{1}{N_1}\sum_{i=1}^{N_1}\sum_{t=1}^{T}q_{it}^{s}q_{it-1}^{s}\right) &= \rho\sigma_e^2 T\underset{N_1\to\infty}{\text{plim}}\left(\frac{1}{N_1}\sum_{i=1}^{N_1}\frac{1}{1-\rho_i^2}\right)\\
&= \rho\sigma_e^2 T\lambda
\end{aligned}
$$

where we use the fact $\text{plim}_{N_1\to\infty}\frac{1}{N_1}\sum_{i=1}^{N_1}\mu_i = 0$ by assumption we made in the above.

$$
\begin{aligned}
\underset{N_2\to\infty}{\text{plim}}\left(\frac{1}{N_2}\sum_{i=1}^{N_2}\sum_{t=1}^{T}\left(q_{it}^{N}\right)^2\right) &= \sigma_e^2\sum_{t=1}^{T}t = \sigma_e^2\frac{T\left(T+1\right)}{2}\\
\underset{N_2\to\infty}{\text{plim}}\left(\frac{1}{N_1}\sum_{i=1}^{N_1}\sum_{t=1}^{T}\left(q_{it}^{s}\right)^2\right) &= \sigma_e^2 T\underset{N_1\to\infty}{\text{plim}}\left(\frac{1}{N_1}\sum_{i=1}^{N_1}\frac{1}{1-\rho_i^2}\right)\\
&= \sigma_e^2 T\lambda
\end{aligned}
$$

Hence we have

$$\begin{aligned}
\underset{N\to\infty}{\text{plim}}\hat{\rho} &= \frac{\pi\rho\sigma_e^2 T\lambda + (1-\pi)\,\sigma_e^2\frac{T(T+1)}{2}}{\pi\sigma_e^2 T\lambda + (1-\pi)\,\sigma_e^2\frac{T(T+1)}{2}} \\
&= \frac{\pi\rho\lambda + (1-\pi)\,\frac{(T+1)}{2}}{\pi\lambda + (1-\pi)\,\frac{(T+1)}{2}}
\end{aligned}$$

Time aggregation bias

Working (1960) assumes that the underlying time series of interest evolves according to the driftless random walk,

$$x_j = x_{j-1} + u_j. \tag{10}$$

Here, $u_j \overset{iid}{\sim} (0,1)$. The intervals at which the observations are reported are indexed by $t = 1,\ldots T$. Within each reporting interval there are M subintervals at which the x_j are observed. The reported observations are period averages at time intervals $j = tM$, for $t = 1,\ldots T$. Denoting the time averaged observations with a tilde, the observable data are

$$\begin{aligned}
\widetilde{x}_t &= \frac{1}{M}\left(x_{(t-1)M} + x_{(t-1)M+1} + \cdots + x_{tM}\right) \\
&= \frac{1}{M}\sum_{j=1}^{M} x_{tM-j}
\end{aligned}$$

For concreteness, if we let $M = 2$, then $\widetilde{x}_t = \frac{1}{2}(x_j + x_{j-1})$, and $\Delta\widetilde{x}_t = \frac{1}{2}(x_j + x_{j-1} - x_{j-2} - x_{j-3}) = \frac{1}{2}(v_j + 2v_{j-1} + v_{j-2})$, $\Delta\widetilde{x}_{t-1} = \frac{1}{2}(x_{j-2} + x_{j-3} - x_{j-4} - x_{j-5}) = \frac{1}{2}(v_{j-2} + 2v_{j-3} + v_{j-4})$. The econometrician studies the time dependence between observations by computing the covariance between period changes in the time averaged observations. The complication is that now both $\Delta\widetilde{x}_{t-1}$ and $\Delta\widetilde{x}_t$ contain v_{j-2}, which gives $\mathrm{E}(\Delta\widetilde{x}_t\Delta\widetilde{x}_{t-1}) = 1/4$. The time averaging has induced artificial serial correlation into the random walk sequence because the truth is $\mathrm{E}(\Delta x_j \Delta x_{j-1}) = 0$. Working shows that as M gets large, the correlation between $\Delta\widetilde{x}_t$ and $\Delta\widetilde{x}_{t-1}$ approaches 1/4. The correlation is 0.235 even when the number of subintervals M is as small as 5

The bias arises as a result of induced endogeneity between u_t and $\widetilde{q}_{t-1}$.The error term u_t follows an MA(1) so that an alternative option to getting a consistent estimate of $\rho = \phi^M$ is to estimate an ARMA(1,1) model to $\widetilde{q}_t$. While it may seem that the bias might vanish as $M \to \infty$,it is inappropriate to take this limit for fixed ϕ, because in applications, we do not observe corresponding reductions in $\widehat{\rho}$ when this is done. Instead, the limit should be taken for a fixed value of ρ. This requires letting $M \to \infty$ simultaneously with $\phi \to 1$ in such a way to keep ρ constant. The nature of the time aggregation bias is

$$\rho = \phi^M < \mathrm{E}(\widehat{\rho}) < \phi$$

To fix ideas, suppose that each time interval has 2 subintervals $M = 2$, from which the underlying observations are averaged. Then, it can be seen that

$$\widetilde{q}_{t+1} = \phi^2 \widetilde{q}_t + \frac{1}{2}(e_4 + (1+\phi)e_3 + \phi e_2)$$

While the coefficient on $\widetilde{q}_t$ declines, $(\phi^2 < \phi)$, the last component e_2 of the composite error term is positively correlated with $\widetilde{q}_t$ which results in an upward bias in the estimator.

Combining point and time-averaged data.

Here, we show that when point-in-time sampled nominal exchange rates are combined with time-averaged price indices that the time-aggregation bias exhibits nuisance parameter dependencies. As a result, it is not possible to obtain meaningful corrections for time-aggregation bias when ρ is estimated using quasi time-averaged observations.

Let s be the log nominal exchange rate and $P = p - p^*$ be the log price differential where s and P follow a permanent-transitory components process. that evolve according to

$$\begin{aligned} s_j &= z_j + u_j^s \\ P_j &= z_j - u_j^P \end{aligned}$$

where $z_j = z_{j-1} + v_j \quad v_j \sim iid(0, \sigma_v^2)$, and

$$\begin{aligned} u_j^s &= \phi u_{j-1}^s + e_j^s, \\ u_j^P &= \phi u_{j-1}^P + e_j^P \end{aligned}$$

where the sum of the transient components follows the AR(1),

$$U_j \equiv u_j^s + u_j^P = \phi U_{j-1} + e_j$$

$e_j \sim iid(0, \sigma_e^2)$. Let Q be the quasi-time averaged real exchange rate and $\widetilde{q}$ be the pure time-averaged real exchange rate. Then the quasi-time averaged rate is,

$$\begin{aligned} Q_{Mt} &= s_{Mt} - \frac{1}{M}\sum_{j=1}^{M} P_{Mt-j} \\ &= \underbrace{z_{Mt} - \frac{1}{M}\sum_{j=1}^{M} z_{Mt-j}}_{(A)} + u_{Mt}^s - \frac{1}{M}\sum_{j=1}^{M} u_{Mt-j}^P \end{aligned} \quad (11)$$

To evaluate the term (A), since $z_{Mt} = z_{M(t-1)} + \sum_{j=1}^{M} v_{Mt-j}$, it follows that

$$
\begin{aligned}
z_{Mt} - \frac{1}{M}\sum_{j=1}^{M} z_{Mt-j} &= z_{M(t-1)} + \sum_{j=1}^{M} v_{tM-j} - z_{M(t-1)} - \frac{1}{M}\sum_{j=1}^{M}\sum_{k=1}^{M-j} v_{Mt-k} \\
&= \sum_{j=1}^{M} v_{tM-j} - \frac{1}{M}\sum_{j=1}^{M}\sum_{k=1}^{M-j} v_{Mt-k} \qquad (12)
\end{aligned}
$$

Substitute (12) into (11) to get

$$
Q_{Mt} = \sum_{j=1}^{M} v_{Mt-j} - \frac{1}{M}\sum_{j=1}^{M}\sum_{k=1}^{M-j} v_{Mt-k} + u^{s}_{Mt} - \frac{1}{M}\sum_{j=1}^{M} u^{P}_{Mt-j} \qquad (13)
$$

From (13), it is seen that Q_{Mt} depends on three innovations, v, u^s and u^P. It follows that the autocorrelation coefficient of Q_{Mt}, will depend on correlation between the two transient components (we assumed above that the innovation to the permanent component is iid). The AR(1) structure of the daily real exchange rate implies an ECM(0) where

$$
\begin{aligned}
\Delta s_j &= \lambda \left(s_{j-1} - p_{j-1}\right) + e^{s}_{j} \\
\Delta p_j &= \left(1 - \lambda - \rho\right)\left(s_{j-1} - p_{j-1}\right) + e^{P}_{j}
\end{aligned}
$$

and

$$
\begin{pmatrix} e^{s}_{j} \\ e^{P}_{j} \end{pmatrix} = iidN\left(0, \begin{bmatrix} 1 & \psi \\ \psi & 1 \end{bmatrix}\right)
$$

To examine the sensitivity of the autocorrelation coefficient to ψ, we conduct a Monte Carlo experiment with 500 replications for $T = 2000$, $M = 12$.We computed the mean values of $\hat{\rho}$ with quasi time aggregated observations well as with 'pure' time aggregated observations. We found that the autocorrelation coefficient ρ can be very sensitive to ψ. For example, let ρ_1 be the autocorrelation coefficient for quasi time averaged observations. Setting $\phi = 0.998$ so that $\phi^{12} = 0.998^{12} = 0.976$ which is similar to our point estimate in applications, we find for $\lambda = 0.05, \psi = -0.8$, $E\left(\hat{\rho}_1 - \phi^M\right) = 0.06$, but for $\lambda = -0.3, \psi = 0.8$, we get $E\left(\hat{\rho}_1 - \phi^M\right) = -0.86$.

Thus, in order to adjust for time-aggregation bias in quasi time-averaged real exchange rates, one would need to have access to the underlying point sampled observations. But if those were available, one would perform direct estimation on the point sampled data and time-aggregation bias would not be an issue.

Combined Nickel and time-aggregation bias in LSDV estimator

We state the bias function $B(\rho, T)$. Under time aggregation, $\rho = \phi^M$. The LSDV estimator has the limit as $N \to \infty$

$$\hat{\rho} = B(\rho, T) = \frac{A_1 - A_2 (T-1)^{-2}}{B_1 - B_2} \tag{14}$$

where

$$\begin{aligned}
A_1 &= (T-1)\,\phi\left(1-\phi^M\right)^2, \\
A_2 &= M(T-2)\left(1-\phi^2\right) + \phi^{M(T-1)}\left[2\phi + \phi\left(1-\phi^M\right)^2\right] - 2\phi^{M+1} \\
B_1 &= M(T-2)\left(1-\phi^2\right) \\
B_2 &= 2\phi\left\{(T-1)\left(1-\phi^M\right) - \frac{1}{T-1}\left(1-\phi^{(T-1)M}\right)\right\}
\end{aligned}$$

Here we provide the derivation for the bias function. The LSDV estimator is

$$\hat{\rho}_{LSDV} = \frac{\sum_{i=1}^N \sum_{t=2}^T q_{it}q_{it-1} - \frac{1}{T-1}\sum_{i=1}^N \left(\sum_{t=2}^T q_{it}\right)\left(\sum_{t=2}^T q_{it-1}\right)}{\sum_{i=1}^N \sum_{t=2}^T q_{it-1}^2 - \frac{1}{T-1}\sum_{i=1}^N \left(\sum_{t=2}^T q_{it-1}\right)^2}$$

Without loss of generality, set $\frac{1}{N}\sum \sigma_i^2 = 1$. As $N \to \infty$,

$$\text{plim}_{N\to\infty}\frac{1}{N}\sum_{i=1}^N \sum_{t=2}^T q_{it}q_{it-1} = (T-1)\sum_{i=1}^M \sum_{j=1}^M \phi^{M+j-i} = (T-1)\,\frac{1}{M}\frac{\phi\left(1-\phi^M\right)^2}{(1-\phi)^2},$$

Note that for any t,

$$\begin{aligned}
Eq_{it}q_{it-1} &= \frac{1}{M^2}E\left(q^+_{i(t-1)M+1} + \cdots + q^+_{itM}\right)\left(q^+_{i(t-2)M+1} + \cdots + q^+_{i(t-1)M}\right) \\
&= \frac{1}{M^2}\sum_{i=1}^M \sum_{j=1}^M \phi^{M+j-i} = \frac{1}{M}\frac{\phi\left(1-\phi^M\right)^2}{(1-\phi)^2},
\end{aligned}$$

$$Eq_{i1}q_{i1+m} = \frac{1}{M^2}\sum_{j=1}^M \phi^{(m-1)M-j+1}\frac{1-\phi^M}{1-\phi} = \frac{1}{M}\frac{\phi^{(m-1)M+1}\left(1-\phi^M\right)^2}{(1-\phi)^2} \quad \text{for } m > 0$$

Let the point-sampled data be denoted by a superscript $+$. Then

$$Eq_{it}^2 = \frac{1}{M^2}E\left(q^+_{i(t-1)M+1} + \cdots + q^+_{itM}\right)^2 = \frac{1}{M}\frac{M\left(1-\phi^2\right) - 2\phi\left(1-\phi^M\right)}{(1-\phi)^2}$$

Hence

$$\begin{aligned}\text{plim}_{N\to\infty}\frac{1}{N}\sum_{i=1}^{N}\sum_{t=2}^{T}q_{it-1}^{2} &= \text{plim}_{N\to\infty}\frac{1}{N}\sum_{i=1}^{N}\left(q_{i1}^{2}+\cdots+q_{i,T-1}^{2}\right) \\ &= (T-1)\frac{1}{M}\frac{M\left(1-\phi^{2}\right)-2\phi\left(1-\phi^{M}\right)}{\left(1-\phi\right)^{2}}, \qquad (15)\end{aligned}$$

To calculate additional terms due to the inclusion of unknown constant, we need to know

$$\begin{aligned}\text{plim}_{N\to\infty}\frac{1}{N}\sum_{i=1}^{N}\left(\sum_{t=2}^{T}q_{it-1}\right)^{2} &= \text{plim}_{N\to\infty}\frac{1}{N}\sum_{i=1}^{N}\frac{1}{M^{2}}\left(q_{i1}^{+}+\cdots+q_{i(T-1)M}^{+}\right)^{2} \\ &= \frac{1}{M}\frac{M\left(T-1\right)\left(1-\phi^{2}\right)-2\phi\left(1-\phi^{(T-1)M}\right)}{\left(1-\phi\right)^{2}} \qquad (16)\end{aligned}$$

and

$$\begin{aligned}&\text{plim}_{N\to\infty}\frac{M}{N}\sum_{i=1}^{N}\left(\sum_{t=2}^{T}q_{it}\right)\left(\sum_{t=2}^{T}q_{it-1}\right) \\ =\ &\text{plim}_{N\to\infty}\frac{M}{N}\sum_{i=1}^{N}\left(\sum_{t=2}^{T}q_{it-1}-q_{i1}+q_{iT}\right)\left(\sum_{t=2}^{T}q_{it-1}\right) \\ =\ &\text{plim}_{N\to\infty}\frac{M}{N}\sum_{i=1}^{N}\left\{\left(\sum_{t=2}^{T}q_{it-1}\right)^{2}-\left(\sum_{t=2}^{T}q_{it-1}\right)\left(q_{i1}-q_{iT}\right)\right\}. \qquad (17)\end{aligned}$$

Note that

$$\begin{aligned}&E\left(\sum_{t=2}^{T}q_{it-1}\right)q_{i1} \\ =\ &E\left(q_{i1}^{2}\right)+Eq_{i1}q_{i2}+...+Eq_{i1}q_{iT-1} \\ =\ &\frac{1}{M}\frac{M\left(1-\phi^{2}\right)-2\phi\left(1-\phi^{M}\right)}{\left(1-\phi\right)^{2}}+\frac{1}{M}\left\{\frac{\phi\left(1-\phi^{M}\right)^{2}}{\left(1-\phi\right)^{2}}+...+\frac{\phi^{M(T-2)+1}\left(1-\phi^{M}\right)^{2}}{\left(1-\phi\right)^{2}}\right\} \\ =\ &\frac{1}{M}\frac{M\left(1-\phi^{2}\right)-2\phi\left(1-\phi^{M}\right)}{\left(1-\phi\right)^{2}}+\frac{1}{M}\frac{\phi\left(1-\phi^{M}\right)^{2}}{\left(1-\phi\right)^{2}}\frac{\left(1-\phi^{M(T-1)}\right)}{1-\phi^{M}},\end{aligned}$$

and

$$
\begin{aligned}
& E\left(\sum_{t=2}^{T} q_{it-1}\right) q_{iT} \\
= \;& Eq_{i1}q_{iT} + ... + Eq_{iT-1}q_{iT} = Eq_{i1}q_{i2} + ... + Eq_{i1}q_{iT} \\
= \;& \frac{1}{M}\frac{\phi\left(1-\phi^{M}\right)^{2}}{\left(1-\phi\right)^{2}}\frac{\left(1-\phi^{MT}\right)}{1-\phi^{M}}
\end{aligned}
$$

Hence we have

$$
\begin{aligned}
& \operatorname{plim}_{N\to\infty}\frac{1}{N}\sum_{i=1}^{N}\left(\sum_{t=2}^{T} q_{it-1}\right)\left(q_{i1}-q_{iT}\right) \\
= \;& \frac{1}{M}\frac{M\left(1-\phi^{2}\right)-2\phi\left(1-\phi^{M}\right)}{\left(1-\phi\right)^{2}} + \frac{1}{M}\frac{\phi\left(1-\phi^{M}\right)^{2}}{\left(1-\phi\right)^{2}}\left\{\frac{\left(1-\phi^{M(T-1)}\right)}{1-\phi^{M}} - \frac{\left(1-\phi^{MT}\right)}{1-\phi^{M}}\right\} \\
= \;& \frac{1}{k}\frac{k\left(1-\phi^{2}\right)-2\phi\left(1-\phi^{k}\right)}{\left(1-\phi\right)^{2}} - \frac{1}{k}\frac{\phi\left(1-\phi^{k}\right)^{2}}{\left(1-\phi\right)^{2}}\phi^{k(T-1)} \qquad (18)
\end{aligned}
$$

Plugging (16), (15) and (18) to (17) yields

$$
\begin{aligned}
& \operatorname{plim}_{N\to\infty}\frac{M}{N}\sum_{i=1}^{N}\left(\sum_{t=2}^{T} q_{it}\right)\left(\sum_{t=2}^{T} q_{it-1}\right) \\
= \;& \frac{1}{\left(1-\phi\right)^{2}}\left\{M\left(T-2\right)\left(1-\phi^{2}\right)+\phi^{M(T-1)}\left[2\phi+\phi\left(1-\phi^{M}\right)^{2}\right]-2\phi^{M+1}\right\}
\end{aligned}
$$

Hence the denominator term in (14) is given by

$$
\begin{aligned}
& \left(T-1\right)\frac{M\left(1-\phi^{2}\right)-2\phi\left(1-\phi^{M}\right)}{\left(1-\phi\right)^{2}} - \frac{1}{T-1}\frac{M\left(T-1\right)\left(1-\phi^{2}\right)-2\phi\left(1-\phi^{(T-1)M}\right)}{\left(1-\phi\right)^{2}} \\
= \;& \left(T-2\right)M\left(1-\phi^{2}\right)-2\phi\left\{\left(T-1\right)\left(1-\phi^{M}\right)-\frac{1}{T-1}\left(1-\phi^{(T-1)M}\right)\right\}
\end{aligned}
$$

while the numerator is

$$
\left(T-1\right)\phi\left(1-\phi^{M}\right)^{2}-\frac{1}{T-1}\left\{M\left(T-2\right)\left(1-\phi^{2}\right)+\phi^{M(T-1)}\left[2\phi+\phi\left(1-\phi^{M}\right)^{2}\right]-2\phi^{M+1}\right\}
$$

That is,

$$
\hat{\rho} = \frac{A_1 - A_2\left(T-1\right)^{-2}}{B_1 - B_2}
$$

Fixed effects GLS

The estimator is fully described in Phillips and Sul (2003). Here, we give only a cursory account. In the absence of time-aggregation, the innovations are governed by the single factor model,

$$e_{it} = \delta_i \theta_t + u_{it}$$

where $\delta_i, i = 1, ...N$, are the factor loadings and θ_t is the single driving factor. The u_{it} are serially and mutually independent. Let $e_t = (e_{1t}, ..., e_{Nt})$, $\delta = (\delta_1, \cdots, \delta_N)$, and $u_t = (u_{1t}, ..., u_{Nt})$. Then $E(e_t e_t') = \Sigma_e = \delta\delta' + \Sigma_u$, where $\Sigma_u = E(u_t u_t')$. The factor loadings can be estimated by iterative method of moments after imposing a normalization for the variance of θ_t. This gives

$$\hat{\Sigma}_\varepsilon = \hat{\delta}\hat{\delta}' + \hat{\Sigma}_u$$

where $\hat{\delta} = \left(\hat{\delta}_1, \cdots, \hat{\delta}_N\right)$ and the diagonal elements of $\hat{\Sigma}_u$ are $\frac{1}{T}\sum_{t=1}^{T} \hat{\varepsilon}_{it}^2$, $\hat{\varepsilon}_{it} = \tilde{q}_{it} - \hat{\rho}_m \tilde{q}_{it-1}$ where $\tilde{q}_{it} = q_{it} - \frac{1}{T}\sum q_{it}$, and $\hat{\rho}_m$ is the mean-unbiased estimator of ρ. Having obtained the estimated error covariance matrix, one can apply feasible GLS to obtain efficient estimates of ρ.

When the observations are time aggregated data, the regression error has an MA(1) structure. In this case, we need one further adjustment because feasible GLS should be based on the long run variance of e_{it} rather than the contemporaneous variance of e_{it}. Since e_{it} follows MA(1), the parametric structure of cross section dependence is now $e_{it} = \eta_{it} + \gamma\eta_{it-1}$, where $\eta_{it} = \delta_i \theta_t + u_{it}$. The long run covariance matrix for e_{it} becomes

$$\Omega_e = E(e_t e_t') + E\left(e_t e_{t-1}'\right) + E(e_{t-1} e_t')$$