
CHI-YOUNG CHOI
NELSON C. MARK
DONGGYU SUL

Unbiased Estimation of the Half-Life to PPP Convergence in Panel Data

Three potential sources of bias introduce complications for panel data estimation of the half-life of purchasing power parity deviations. They are bias induced by inappropriate cross-sectional aggregation of heterogeneous coefficients, small-sample estimation bias of dynamic lag coefficients, and bias induced by time aggregation of commodity prices. All of these biases have been addressed individually in the literature. In this paper, we address all three biases in arriving at our estimates. Using an annual panel data set of CPI-based real exchange rates for 21 OECD countries from 1973 to 1998, we obtain a point estimate of the unbiased half-life of 3.0 years with a 95% confidence interval of 2.3–4.2 years.

JEL codes: C32, F31, F47

Keywords: PPP, time aggregation, bias, half-life.

WE ARE INTERESTED in obtaining accurate measurements of the convergence rate to purchasing power parity (PPP) because of its role in informing theoretical work on the role of nominal rigidities and on the relative importance of nominal and real shocks in international macro models. The motivation for using panel data to estimate the convergence half-life of 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 Estimation accuracy is

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 consensus 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 (2004), 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.

We thank David Parsley and an anonymous referee for useful comments that helped to improve the paper.

CHI-YOUNG CHOI is affiliated with the Department of Economics, University of New Hampshire. NELSON C. MARK is affiliated with the Department of Economics and Econometrics, University of Notre Dame and NBER (E-mail: nmark@nd.edu). DONGGYU SUL is affiliated with the Department of Economics, University of Auckland.

Received July 2, 2004; and accepted in revised form November 22, 2004.

Journal of Money, Credit, and Banking, Vol. 38, No. 4 (June 2006)

Copyright 2006 by The Ohio State University

especially important due to the nonlinearity of the half-life formula because small differences in the estimated value of dynamic lag coefficients in the real exchange rate process 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 subject to three potential sources of bias. These biases have been addressed individually in the literature. In this short paper, we address all three potential sources of bias to arrive at a final and unbiased estimate of the half-life. Using an annual sample of 21 OECD country CPI-based real exchange rates from 1973 to 1998, and controlling for multiple sources of bias, we estimate the half-life to be 3 years with a 95% confidence interval ranging from 2.3 to 4.2 years. This approximately brings us back to point estimates obtained by the uncorrected least squares dummy variable method and lies at the short end of Rogoff's (1996) consensus half-life estimate of 3–5 years.

The first potential source of bias that we address is inappropriate pooling across cross-sectional units. If the real exchange rates of different countries exhibit heterogeneous rates of convergence to PPP, then the panel data estimator of a common autoregressive coefficient can be biased upwards and the data should not be pooled. Imbs et al. (2005) study how sectoral heterogeneity in convergence rates to the law of one price can lead to upward bias in the estimated half-life. Chen and Engel (2005), on the other hand, find that sectoral heterogeneity is not a quantitatively important source of bias. We do not directly address sectoral heterogeneity of convergence rates in this paper, but we allow for the possibility that sectoral heterogeneity can induce heterogeneity in country-level data.

Second, we address downward small-sample estimation bias that results when the dynamic regression is run with a constant. 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 the downward bias can be seen by noting that least squares estimation of an autoregression with a constant is equivalent to running the regression without a constant on observations that are deviations from the sample mean. The problem then is that, for any observation, the regression error is correlated with current and future values of the dependent variable which are embedded in the sample mean and which in turn is a component of the independent variable. This induced correlation between the regression error and the sample mean creates the downward bias. In fixed-effects estimation with panel data, 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. We henceforth refer to the bias in the panel data context as "Nickell bias."²

Third, we address the upward bias that results because 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 aggregation of the data were first discussed

2. The LSDV method is pooled OLS with fixed effects. See Hsiao (2003).

by Working (1960). Taylor (2001) extends Working's analysis to the study of PPP and shows that time aggregation leads to an upward bias in the estimated half-life.

The remainder of the paper is organized as follows. The next section discusses half-life measurement. Section 2 discusses the three potential biases that we examine. Section 3 outlines our bias-adjustment strategies and presents the empirical results. Section 4 concludes. The Appendix contains derivations for many of the results presented in the text.

1. HALF-LIFE MEASUREMENT

Let the real exchange rate for country $i = 1, \dots, N$ evolve according to a first-order autoregression (AR(1)), $q_{it} = \alpha_i + \rho q_{it-1} + e_{it}$, where e_{it} is a mean-zero serially uncorrelated innovation with finite fourth moments. 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 divergence from PPP to dissipate by one half. In the AR(1) case, it is t^* such that $E(q_{it^*}) = e_1/2 = 1/2$, which takes the convenient form

$$t^* = H(\rho) = \frac{\ln(0.5)}{\ln(\rho)}. \quad (1)$$

Due to the nonlinear nature of $H(\rho)$, small variations in ρ in the region near unity lead to disproportionately large variations in the half-life.³ Thus, if the estimator of ρ is biased, failure to adjust for the bias can produce substantively misleading estimates of the half-life.

For more complicated dynamic models that have additional lags or moving average error terms, Equation (1) gives an approximation to the true half-life. For these models, the half-life can also be computed by impulse response analysis but a knotty problem associated with this approach is that the half-life may not be unique on account of nonmonotonic impulse responses. However, we are able to obtain relatively clean and straightforward results by employing annual data. The AR(1) specification is appropriate for annual data, whereas adequate modeling of monthly or quarterly real exchange rates requires higher-ordered AR specifications.⁴

2. THREE POTENTIAL SOURCES OF BIAS

This section reviews three potential sources of bias discussed in the literature. Section 3 presents our strategy for accounting for these biases.

3. For example, $H(0.93) = 9.56$, $H(0.95) = 13.5$, $H(0.97) = 22.8$.

4. Since the annual observations are built up from the quarterly (monthly) data, the estimated half-life should be robust to the choice of data if the biases are properly controlled for. This is in fact what we find. However, using annual data, we are able to avoid complications arising from nonuniqueness of the half-life. We also avoid lag-length selection of the autoregression which turns out not to be innocuous. The analysis with monthly or quarterly data is substantially less straightforward than the analysis on annual data. Following the principle of Occam's razor, we center our analysis on annual data. The results from monthly and quarterly data are contained in a separate appendix that is available from the authors upon request.

2.1 Cross-sectional Aggregation Bias

Imbs et al. (2005) study how sectoral heterogeneity in the convergence rate to the law of one price across commodity categories in the price level can induce an upward bias in the estimated half-life of PPP deviations. Chen and Engel (2005), on the other hand, conclude that sectoral heterogeneity is not a quantitatively important source of bias. We do not directly address the issue of sectoral heterogeneity, but we do allow for the possibility that sectoral heterogeneity might induce heterogeneity in the convergence rate to PPP in aggregate country-level real exchange rates.

To see how cross-sectional heterogeneity can bias the panel estimator, suppose that the real exchange rate for country i follows⁵

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

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 Equation (3) into Equation (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} . Decomposing the pooled OLS estimator gives

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

The term $A(N, T)$ is standard. The bias introduced by aggregating across heterogeneous cross-sectional coefficients is $B(N, T)$. This bias is not likely to be quantitatively important if all of the country real exchange rates are covariance stationary and the distribution of v_i is symmetric because the average of the terms $v_i (\sum_{t=1}^T q_{it-1}^2)$ will be close to zero. The bias is potentially important 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 pooled OLS estimator can be expressed as

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

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

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

6. See the Appendix for the derivation.

2.2 Nickell Bias

Consider estimating an autoregression, but instead of including a constant, it is estimated with observations that are deviations from the sample mean. Then for any observation t , the regression error is correlated with current and future values of the dependent variable. Because these future values are embedded in the sample mean which is now a component of the explanatory variable, the least squares estimator is biased when a constant is included in the regression. The use of panel data does not eliminate this small-sample bias.

Nickell (1981) studied the properties of the LSDV estimator for the dynamic panel regression model with fixed effects for cross-sectionally independent observations. 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 $N \rightarrow \infty$). For the LSDV estimator in the panel AR(1) model with fixed effects, Nickell shows

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

which is biased downwards.

2.3 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, then time-averaging this process induces a moving average error into the reported (time-averaged) first differences. The analyst who 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 analysis to the case where the true point-sampled process follows a stationary AR(1). In the PPP context, 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. Taylor 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

7. Tiao (1972) and Brewer (1973) also develop econometric implications of time aggregation. Rossana and Seater (1992), Marcellino (1999), and Breitung and Swanson (2002) study the effects of time aggregation on exogeneity tests and forecasting.

the underlying price process is observable (if the true point-sampled process is observable daily and is averaged to form annual data, then $M = 365$). The annual observations are reported as period averages at the annual time intervals $j = M, 2M, \dots, TM$. Assuming that the dynamics of the underlying point-in-time daily real exchange rate process evolves according to the AR(1) process $q_{ij} = \alpha_i + \phi q_{ij-1} + e_{ij}$, 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 subintervals, 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

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

which overstates the true 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, this is an inappropriate strategy because 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.

3. UNBIASED HALF-LIFE ESTIMATION

3.1 The Data

Our data are annual real exchange rates of 21 OECD countries which we construct by combining annual nominal exchange rates and annual consumer price indices from 1973 to 1998 which results in 25 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 (2004) 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.

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

3.2 Cross-sectional Heterogeneity

Here, 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 the dynamic regressions associated with each of the numeraire countries and conduct homogeneity tests on recursive mean-adjusted seemingly unrelated regression estimates of ρ .⁹ The results of the homogeneity test are reported in Table 1. Homogeneity is rejected at the 5% level only when Germany serves as the numeraire country. Because the evidence against homogeneity in this data set is quite weak, cross-country heterogeneity in the autoregressive coefficient does not appear to be a significant source of bias. We proceed by assuming that pooling is appropriate.

3.3 Combined Nickell and Time Aggregation Bias Adjustments

In the absence of aggregation bias, the adjustment for Nickell bias can be done directly by panel mean unbiased estimation. This can be obtained by estimating ρ by LSDV and then using the inverse function of the bias formula to obtain the mean unbiased estimator $\hat{\rho}_{\text{MUE}} = m^{-1}(\hat{\rho}_{\text{LSDV}})$, where $\hat{\rho}_{\text{LSDV}}$ is the LSDV estimator and the function $m(\cdot)$ is given in Equation (7).¹⁰ When the data are time-averaged, there is an interaction between the Nickell bias and the time aggregation bias which necessitates a further adjustment in the mean function for panel mean unbiased estimation. An analytical characterization of the combined bias is provided in the Appendix. We denote the formula that simultaneously accounts for both sources of bias as $B(\rho, M, T)$.

TABLE 1
HOMOGENEITY TEST (REAL EXCHANGE RATES IN 21 OECD COUNTRIES 1973–1998)

Nomeraire country	Wald statistic	P-value	Nomeraire country	Wald statistic	P-value
Australia	11.223	0.916	Japan	18.151	0.512
Austria	30.099	0.051	Netherlands	23.935	0.199
Belgium	24.442	0.180	New Zealand	21.007	0.336
Canada	18.889	0.464	Norway	8.970	0.974
Denmark	18.977	0.458	Portugal	21.876	0.290
Finland	5.794	0.998	Spain	6.908	0.995
France	20.079	0.390	Sweden	12.142	0.879
Germany	31.251	0.038	Switzerland	15.529	0.688
Greece	9.560	0.963	U.K.	19.669	0.415
Ireland	7.345	0.992	U.S.A.	13.948	0.787
Italy	3.465	1.000			

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

10. Murray and Papell (2002) used median unbiased estimation in a univariate PPP analysis. Murray and Papell (2004) employed median unbiased estimation to a panel PPP study. In the univariate context, the mean and median unbiased estimators are not equivalent in small samples. In the panel context, however, Phillips and Sul (2004) show that even with a moderate cross-sectional dimension (e.g., $N = 5$) the difference between the mean and median unbiased estimators is trivial.

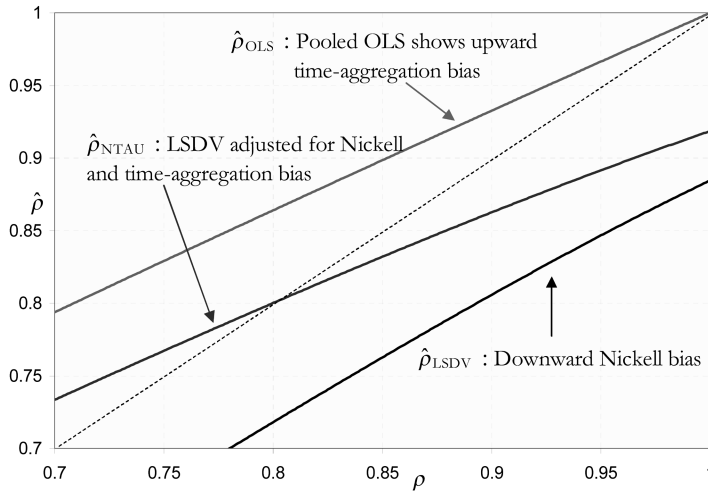


FIG. 1. Biases of Pooled Estimators with Pure Time-aggregated Data ($T = 25$)

The pure Nickell bias and the time aggregation bias work in opposite directions. A decomposition of the opposing bias factors is shown in Figure 1. Here, the true value of ρ is plotted on the horizontal axis, and the LSDV probability limit is plotted on the vertical axis.¹¹ The top line shows the effect of time aggregation in panel data. This is the probability limit of the pooled OLS estimator on time-aggregated data with $M = 12$ and no regression constant. Here, correcting a pooled OLS point estimate of 0.9 (implied half-life of 6.6 years) for time aggregation bias yields an adjusted value of 0.85 for ρ (implied half-life of 4.3 years). As $\rho \rightarrow 1$, the upward time aggregation bias vanishes. The bottom line shows the effect of pure Nickell bias which is the LSDV probability limit from Equation (7). Here, an LSDV point estimate of 0.9 has a mean unbiased value of 0.95 for ρ (implied half-life of 13.5 years). The center line plots $B(\rho, M, T)$ which 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 a net upward (downward) combined bias.

Thus, a strategy to simultaneously correct for Nickell and time aggregation bias is to first estimate ρ by LSDV and then use the inverse function to obtain a Nickell and time aggregation unbiased estimate,¹²

$$\hat{\rho}_{\text{NTAU}} = B^{-1}(\hat{\rho}_{\text{LSDV}}, M, T). \tag{9}$$

To this proposed correction, we make one additional adjustment. Because LSDV does not exploit the cross-sectional covariance structure of the observations in

11. The probability limits are for $N \rightarrow \infty$ but for fixed $T = 25$ which corresponds to the number of time series observations in our data set.

12. NTAU stands for the Nickell and Time Aggregation Unbiased estimator.

TABLE 2
 PANEL FEASIBLE GLS ESTIMATION

Numeraire	No bias corrections		Nickell bias corrected		Time aggregation bias corrected		Nickell and time aggregation bias corrected			
	$\hat{\rho}$	$H_{0.5}$	$\hat{\rho}$	$H_{0.5}$	$\hat{\rho}$	$H_{0.5}$	$\hat{\rho}_{GNTAU}$	$H_{0.025}$	$H_{0.5}$	$H_{0.975}$
Australia	0.820	3.5	0.893	6.1	0.717	2.1	0.800	2.2	3.1	4.8
Austria	0.884	5.6	0.933	10.0	0.766	2.6	0.855	3.2	4.4	6.9
Belgium	0.852	4.3	0.930	9.6	0.760	2.5	0.848	2.9	4.2	7.2
Canada	0.827	3.6	0.870	5.0	0.692	1.9	0.772	2.0	2.7	3.8
Denmark	0.829	3.7	0.909	7.3	0.739	2.3	0.824	2.6	3.6	5.4
Finland	0.797	3.1	0.864	4.7	0.684	1.8	0.766	2.0	2.6	3.7
France	0.813	3.3	0.887	5.8	0.712	2.0	0.795	2.3	3.0	4.1
Germany	0.852	4.3	0.926	9.0	0.758	2.5	0.845	3.0	4.1	6.3
Greece	0.815	3.4	0.887	5.8	0.709	2.0	0.795	2.3	3.0	4.3
Ireland	0.809	3.3	0.858	4.5	0.674	1.8	0.755	1.9	2.5	3.5
Italy	0.808	3.3	0.874	5.1	0.694	1.9	0.777	2.1	2.8	3.8
Japan	0.827	3.6	0.893	6.1	0.720	2.1	0.803	2.3	3.2	4.9
Netherlands	0.875	5.2	0.933	10.0	0.769	2.6	0.855	3.1	4.4	7.2
New Zealand	0.811	3.3	0.883	5.6	0.709	2.0	0.792	2.3	3.0	4.2
Norway	0.810	3.3	0.858	4.5	0.674	1.8	0.755	1.9	2.5	3.3
Portugal	0.814	3.4	0.867	4.9	0.687	1.8	0.766	2.0	2.6	3.7
Spain	0.807	3.2	0.887	5.8	0.709	2.0	0.792	2.2	3.0	4.2
Sweden	0.813	3.3	0.858	4.5	0.677	1.8	0.755	1.9	2.5	3.4
Switzerland	0.821	3.5	0.887	5.8	0.709	2.0	0.792	2.3	3.0	4.1
U.K.	0.823	3.6	0.880	5.4	0.704	2.0	0.786	2.1	2.9	4.2
U.S.A.	0.825	3.6	0.896	6.3	0.723	2.1	0.806	2.3	3.2	5.1
Median	0.820	3.5	0.887	5.8	0.709	2.0	0.792	2.3	3.0	4.2

NOTE: $H_{0.025}$, $H_{0.5}$, and $H_{0.975}$ are the 2.5, 50, and 97.5 percentiles of the half-life distribution.

estimation, we achieve an improvement in efficiency by using a panel generalized least squares (GLS) estimator with fixed effects. When the cross-sectional dependence has a single-factor structure, the Nickell bias of the fixed effects GLS estimator is independent of both the factor loadings and the unobserved factor (Phillips and Sul 2004). This independence allows us to apply the mean adjustment in Equation (9) with the panel GLS estimator in place of $\hat{\rho}_{LSDV}$. Call the estimator

$$\hat{\rho}_{GNTAU} = B^{-1}(\hat{\rho}_{GLS}, M, T).$$

A description of the estimator is given in the Appendix.

As there is no definitive account of the exact number of subintervals over which consumer prices are time averaged, we performed estimation for $M = 12, 130, 260,$ and 365 and found that the results are robust to these variations. As a result, we report the results only for $M = 12$.¹³ Table 2 reports feasible panel GLS estimates and associated half-lives obtained under corrections for the various biases. The median half-life obtained from uncorrected GLS estimates is 3.5 years. Our median point estimate after adjusting only for Nickell bias is 0.89 which yields a half-life of 5.8 years. The median half-life obtained after correcting only for time aggregation bias but not for Nickell bias is 2 years. Our median estimate of

13. The results for the alternative values of M are available from the authors upon request.

$\hat{\rho}_{\text{GNTAU}} = B^{-1}(\hat{\rho}_{\text{GLS}}, 12, 25)$ which simultaneously corrects for Nickell bias and time aggregation bias is 0.79 with an associated half-life of 3 years. As a point of comparison, we note that our results are consistent with those of Murray and Papell (2004) who employ quarterly data. They obtain a median unbiased half-life of 4.6 years when the U.S.A. is used as the numeraire when the lag length of the dynamic real exchange rate regressions by Ng and Perron's (2001) modified AIC procedure. Their estimate is somewhat above our estimate of 3.2 years after correcting for both Nickell bias and time aggregation bias (U.S.A. numeraire) but is a bit lower than our Nickell bias corrected half-life estimate of 6.3 years.

4. CONCLUSION

PPP research, desperate for larger sample sizes to improve precision and confidence in empirical estimates, has turned to the analysis of panel data. Three potential sources of bias in the estimation of the half-life to PPP convergence have been discussed in the literature. We found that in country-level data, cross-sectional heterogeneity of convergence rates to PPP do not appear to be a quantitatively important source of bias. The remaining two sources are the downward bias of the panel fixed-effects estimator and time aggregation bias. Simultaneously controlling for these latter two sources of bias yields a point estimate of 3 years for the half-life with reasonably tight confidence intervals.

APPENDIX

Derivation of Equation (6)

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

$$\begin{aligned} & \pi \text{plim}_{N_1 \rightarrow \infty} \left(\frac{1}{N_1} \sum_{i=1}^{N_1} \sum_{t=1}^T q_{it}^s q_{it-1}^s \right) + (1 - \pi) \text{plim}_{N_2 \rightarrow \infty} \left(\frac{1}{N_2} \sum_{i=1}^{N_2} \sum_{t=1}^T q_{it}^N q_{it-1}^N \right) \\ &= \frac{\pi \text{plim}_{N_1 \rightarrow \infty} \left(\frac{1}{N_1} \sum_{i=1}^{N_1} \sum_{t=1}^T (q_{it-1}^s)^2 \right) + (1 - \pi) \text{plim}_{N_2 \rightarrow \infty} \left(\frac{1}{N_2} \sum_{i=1}^{N_2} \sum_{t=1}^T (q_{it-1}^N)^2 \right)}{\pi \text{plim}_{N_1 \rightarrow \infty} \left(\frac{1}{N_1} \sum_{i=1}^{N_1} \sum_{t=1}^T q_{it}^s q_{it-1}^s \right) + (1 - \pi) \text{plim}_{N_2 \rightarrow \infty} \left(\frac{1}{N_2} \sum_{i=1}^{N_2} \sum_{t=1}^T q_{it}^N q_{it-1}^N \right)}. \end{aligned}$$

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

$$\text{plim}_{N_2 \rightarrow \infty} \left(\frac{1}{N_2} \sum_{i=1}^{N_2} \sum_{t=1}^T q_{it}^N q_{it-1}^N \right) = \sigma_e^2 \sum_{t=1}^T t = \sigma_e^2 \frac{T(T+1)}{2},$$

$$\text{plim}_{N_1 \rightarrow \infty} \left(\frac{1}{N_1} \sum_{i=1}^{N_1} \sum_{t=1}^T q_{it}^s q_{it-1}^s \right) = \rho \sigma_e^2 T \text{plim}_{N_1 \rightarrow \infty} \left(\frac{1}{N_1} \sum_{i=1}^{N_1} \frac{1}{1 - \rho_i^2} \right) = \rho \sigma_e^2 T \lambda,$$

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

$$\text{plim}_{N_2 \rightarrow \infty} \left(\frac{1}{N_2} \sum_{i=1}^{N_2} \sum_{t=1}^T (q_{it}^N)^2 \right) = \sigma_e^2 \sum_{t=1}^T t = \sigma_e^2 \frac{T(T+1)}{2},$$

$$\text{plim}_{N_1 \rightarrow \infty} \left(\frac{1}{N_1} \sum_{i=1}^{N_1} \sum_{t=1}^T (q_{it}^s)^2 \right) = \sigma_e^2 T \text{plim}_{N_1 \rightarrow \infty} \left(\frac{1}{N_1} \sum_{i=1}^{N_1} \frac{1}{1 - \rho_i^2} \right) = \sigma_e^2 T \lambda.$$

Hence, we have

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

Time Aggregation Bias

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

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

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

$$\begin{aligned} \tilde{x}_t &= \frac{1}{M} (x_{(t-1)M} + x_{(t-1)M+1} + \dots + x_{tM}) \\ &= \frac{1}{M} \sum_{j=1}^M x_{tM-j} . \end{aligned}$$

For concreteness, if we let $M = 2$, then $\tilde{x}_t = 1/2(x_j + x_{j-1})$, $\Delta\tilde{x}_t = 1/2(x_j + x_{j-1} - x_{j-2} - x_{j-3}) = 1/2(v_j + 2v_{j-1} + v_{j-2})$, and $\Delta\tilde{x}_{t-1} = 1/2(x_{j-2} + x_{j-3} - x_{j-4} - x_{j-5}) = 1/2(v_{j-2} + 2v_{j-3} + v_{j-4})$. 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 $\Delta\tilde{x}_{t-1}$ and $\Delta\tilde{x}_t$ contain v_{j-2} , which gives $E(\Delta\tilde{x}_t\Delta\tilde{x}_{t-1}) = 1/4$. The time averaging has induced artificial serial correlation into the random walk sequence because the truth is $E(\Delta x_j\Delta x_{j-1}) = 0$. Working shows that as M gets large, the correlation between $\Delta\tilde{x}_t$ and $\Delta\tilde{x}_{t-1}$ 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 v_t and \tilde{q}_{t-1} . The error term v_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 \tilde{q}_t . While it may seem that the bias might vanish as $M \rightarrow \infty$, it is inappropriate to take this limit for fixed ϕ , because in applications, we do not observe corresponding reductions in $\hat{\rho}$ when this is done. Instead, the limit should be taken for a fixed value of ρ . This requires letting $M \rightarrow \infty$ simultaneously with $\phi \rightarrow 1$ in such a way to keep ρ constant. The nature of the time aggregation bias is

$$\rho = \phi^M < E(\hat{\rho}) < \phi .$$

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

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

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

Combining Point and Time-averaged Data

Here, we show that the time aggregation bias exhibits nuisance parameter dependencies when point-in-time sampled nominal exchange rates are combined with time-averaged price indices. 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 evolves according to

$$s_j = z_j + u_j^s,$$

$$P_j = z_j - u_j^P,$$

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

$$u_j^s = \phi u_{j-1}^s + e_j^s,$$

$$u_j^P = \phi u_{j-1}^P + e_j^P,$$

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

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

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

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

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

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

Substitute Equation (12) to Equation (11) to get

$$Q_{Mt} = \sum_{j=1}^M v_{Mt-j} - \frac{1}{M} \sum_{j=1}^M \sum_{k=1}^{M-j} v_{Mt-k} + u_{Mt}^s - \frac{1}{M} \sum_{j=1}^M u_{Mt-j}^P. \tag{13}$$

From Equation (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

$$\Delta s_j = \lambda(s_{j-1} - p_{j-1}) + e_j^s,$$

$$\Delta p_j = (1 - \lambda - \rho)(s_{j-1} - p_{j-1}) + e_j^P,$$

and

$$\begin{pmatrix} e_j^s \\ e_j^P \end{pmatrix} = \text{iid } N\left(0, \begin{bmatrix} 1 & \Psi \\ \Psi & 1 \end{bmatrix}\right).$$

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 as 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 $\lambda = 0.05, \psi = -0.8, E(\hat{\rho}_1 - \phi^M) = 0.06$, but for $\lambda = -0.3, \psi = 0.8$, we get $E(\hat{\rho}_1 - \phi^M) = -0.86$.

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 these were available, one would perform direct estimation on the point-sampled data and time aggregation bias would not be an issue.

Combined Nickell 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 $N \rightarrow \infty$

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

where

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

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

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

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

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

Note that for any t ,

$$\begin{aligned} E q_{it} q_{it-1} &= \frac{1}{M^2} E(q_{i(t-1)M+1}^+ + \dots + q_{iM}^+)(q_{i(t-2)M+1}^+ + \dots + q_{i(t-1)M}^+) \\ &= \frac{1}{M^2} \sum_{i=1}^M \sum_{j=1}^M \phi^{M+j-i} = \frac{1}{M} \frac{\phi(1 - \phi^M)^2}{(1 - \phi)^2}, \end{aligned}$$

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

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

$$Eq_{it}^2 = \frac{1}{M^2} E(q_{i(t-1)M+1}^+ + \dots + q_{itM}^+)^2 = \frac{1}{M} \frac{M(1 - \phi^2) - 2\phi(1 - \phi^M)}{(1 - \phi)^2}.$$

Hence,

$$\begin{aligned} \text{plim}_{N \rightarrow \infty} \frac{1}{N} \sum_{i=1}^N \sum_{t=2}^T q_{it-1}^2 &= \text{plim}_{N \rightarrow \infty} \frac{1}{N} \sum_{i=1}^N (q_{i1}^2 + \dots + q_{iT-1}^2) \\ &= (T - 1) \frac{1}{M} \frac{M(1 - \phi^2) - 2\phi(1 - \phi^M)}{(1 - \phi)^2}. \end{aligned} \tag{15}$$

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

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

and

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

Note that

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

and

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

Hence, we have

$$\begin{aligned}
 \text{plim}_{N \rightarrow \infty} \frac{1}{N} \sum_{i=1}^N \left(\sum_{t=2}^T q_{it-1} \right) (q_{i1} - q_{iT}) &= \frac{1}{M} \frac{M(1 - \phi^2) - 2\phi(1 - \phi^M)}{(1 - \phi)^2} \\
 &\quad + \frac{1}{M} \frac{\phi(1 - \phi^M)^2}{(1 - \phi)^2} \left[\frac{(1 - \phi^{M(T-1)})}{1 - \phi^M} - \frac{(1 - \phi^{MT})}{1 - \phi^M} \right] \\
 &= \frac{1}{k} \frac{k(1 - \phi^2) - 2\phi(1 - \phi^k)}{(1 - \phi)^2} \tag{18} \\
 &\quad - \frac{1}{k} \frac{\phi(1 - \phi^k)^2}{(1 - \phi)^2} \phi^{k(T-1)}.
 \end{aligned}$$

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

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

Hence, the denominator term in (14) is given by

$$\begin{aligned}
 (T - 1) \frac{M(1 - \phi^2) - 2\phi(1 - \phi^M)}{(1 - \phi)^2} - \frac{1}{T - 1} \frac{M(T - 1)(1 - \phi^2) - 2\phi(1 - \phi^{(T-1)M})}{(1 - \phi)^2} \\
 = (T - 2)M(1 - \phi^2) - 2\phi\{(T - 1)(1 - \phi^M) - \frac{1}{T - 1}(1 - \phi^{(T-1)M})\},
 \end{aligned}$$

while the numerator is

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

That is,

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

Fixed-effects GLS

The estimator is fully described in Phillips and Sul (2004). 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 $\delta_i, i = 1, \dots, N$, are the factor loadings and θ_t is the single driving factor. The u_{it} are serially and mutually independent. Let $e_t = (e_{1t}, \dots, e_{Nt})$, $\delta = (\delta_1, \dots, \delta_N)$, and $u_t = (u_{1t}, \dots, u_{Nt})$. 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

$$\hat{\Sigma}_e = \hat{\delta} \hat{\delta}' + \hat{\Sigma}_u,$$

where $\hat{\delta} = (\hat{\delta}_1, \dots, \hat{\delta}_N)$ and the diagonal elements of $\hat{\Sigma}_u$ are $\frac{1}{T} \sum_{t=1}^T \hat{\epsilon}_{it}^2$, $\hat{\epsilon}_{it} = \tilde{q}_{it} - \hat{\rho}_m \tilde{q}_{it-1}$, 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

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

LITERATURE CITED

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.

Brewer, Kenneth R.W. (1973). "Some Consequences of Temporal Aggregation and Systematic Sampling for ARMA and ARMAX Models." *Journal of Econometrics* 1, 133–154.

Chen, Shiu-Sheng, and Charles Engel (2005). "Does 'Aggregation Bias' Explain the PPP Puzzle?" *Pacific Economic Review* 10, 49–72.

Chiu, Ru-Lin (2002). "Testing the Purchasing Power Parity in Panel Data." *International Review of Economics and Finance* 11, 349–362.

Choi, Chi-Young (2004). "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* 23, 1159–1186.

Choi, Chi-Young, Nelson C. Mark, and Donggyu Sul (2004). "Bias Reduction by Recursive Mean Adjustment in Panel Data." Manuscript, University of Auckland.

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.

- 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.
- 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, 209–224.
- Hsiao, Cheng (2003). *Analysis of Panel Data*, 2nd edition. New York: Cambridge University Press.
- Imbs, Jean, Haroon Mumtaz, Morten O. Ravn, and Hélène Rey (2005). "PPP Strikes Back: Aggregation and the Real Exchange Rate." *Quarterly Journal of Economics* 120, 1–43.
- Kendall, M.G. (1954). "Note on Bias in the Estimation of Autocorrelation." *Biometrika* 41, 403–404.
- 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.
- 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.
- Marcellino, Massimiliano (1999). "Some Consequences of Temporal Aggregation in Empirical Analysis." *Journal of Business and Economic Statistics* 17, 129–136.
- Marriott, F.H.C., and J.A. Pope (1954). "Bias in the Estimation of Autocorrelations." *Biometrika* 41, 390–402.
- Murray, Christian J., and David H. Papell (2002). "The Purchasing Power Parity Persistence Paradigm." *Journal of International Economics* 56, 1–19.
- Murray, Christian J., and David H. Papell (2004). "Do Panels Help Solve the Purchasing Power Parity Puzzle?" Manuscript, University of Houston.
- Ng, Serena, and Pierre Perron (2001). "Lag Length Selection and the Construction of Unit Root Tests with Good Size and Power." *Econometrica* 69, 1519–1554.
- Nickell, Stephen (1981). "Biases in Dynamic Models with Fixed Effects." *Econometrica* 49, 1417–1426.
- Papell, David H. (2004). "The Panel Purchasing Power Parity Puzzle." Manuscript, University of Houston.
- 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.
- 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.
- Phillips, Peter C.B., and Donggyu Sul (2004). "Bias in Dynamic Panel Estimation with Fixed Effects, Incidental Trends and Cross Section Dependence." Manuscript, University of Auckland.
- Rogoff, Kenneth (1996). "The Purchasing Power Parity Puzzle." *Journal of Economic Literature* 34, 647–668.
- Rossana, Robert J., and John J. Seater (1992). "Aggregation, Unit Roots and the Time-Series Structure of Manufacturing Real Wages." *International Economic Review* 33, 159–179.
- Taylor, Alan M. (2001). "Potential Pitfalls for the PPP Puzzle? Sampling and Specification Biases in Mean-reversion Tests of the Law of One Price." *Econometrica* 69, 473–498.
- Tiao, George C. (1972). "Asymptotic Behavior of Temporal Aggregates of Time Series." *Biometrika* 59, 525–531.
- Working, Holbrook (1960). "Note on the Correlation of First Differences of Averages in a Random Chain." *Econometrica* 28, 916–918.