NBER WORKING PAPER SERIES

DOES "AGGREGATION BIAS" EXPLAIN THE PPP PUZZLE?

Shiu-Sheng Chen
Charles Engel

Working Paper 10304
http://www.nber.org/papers/w10304

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

Does “Aggregation Bias” Explain the PPP Puzzle?
Shiu-Sheng Chen and Charles Engel
NBER Working Paper No. 10304
February 2004
JEL No. F31, C22, C23

ABSTRACT

Recently, Imbs et. al. (2002) have claimed that much of the purchasing power parity puzzle can be explained by “aggregation bias”. This paper re-examines aggregation bias. First, it clarifies the meaning of aggregation bias and its applicability to the PPP puzzle. Second, the size of the “bias” is shown to be much smaller than the simulations in Imbs et. al. (2002) suggest, if we rule out explosive roots in the simulations. Third, we show that the presence of non-persistent measurement error – especially in the Imbs et. al. (2002) data – can make price series appear less persistent than they really are. Finally, it is now standard to recognize that small-sample bias plagues estimates of speeds of convergence of PPP. After correcting small sample bias by methods proposed by Kilian (1998) and by So and Shin (1999), the half-life estimates indicate that heterogeneity and aggregation bias do not help to solve the PPP puzzle.

Shiu-Sheng Chen
Department of Economics
University of Wisconsin
1180 Observatory Drive
Madison, WI 53706-1393
schen9@wisc.edu

Charles Engel
Department of Economics
University of Wisconsin
1180 Observatory Drive
Madison, WI 53706-1393
and NBER
cengel@ssc.wisc.edu

1. Introduction

Rogoff (1996) defines the "purchasing power parity puzzle" as:

How can one reconcile the enormous short-term volatility of real exchange rates with the extremely slow rate at which shocks appear to damp out? Most explanations of short-term exchange rate volatility point to financial factors such as changes in portfolio preferences, short-term asset price bubbles, and monetary shocks. Such shocks can have substantial effects on the real economy in the presence of sticky nominal wages and prices. Consensus estimates for the rate at which PPP deviations damp, however, suggest a half-life of three to five years, seemingly far too long to be explained by nominal rigidities. It is not difficult to rationalize slow adjustment if real shocks – shocks to tastes and technology – are predominant. But existing models based on real shocks cannot account for short-term exchange-rate volatility. (pp. 647-648.)

Algebraically, it is helpful to express this in the following terms. Take a log-linear approximation of the home country and foreign country consumer price indexes:

$$(1) \qquad p_t = \alpha_1 p_{1t} + \alpha_2 p_{2t} + \ldots + \alpha_N p_{Nt} \qquad\qquad p_t^* = \alpha_1 p_{1t}^* + \alpha_2 p_{2t}^* + \ldots + \alpha_N p_{Nt}^* .$$

p_t is the log of the home CPI, and p_t^* the log of the foreign CPI, assumed for convenience to be expressed in a common currency. p_{it} (p_{it}^*) is the price of good *i* in the home (foreign) country. We will assume that the CPI weights, α_i, are the same at home and abroad. The α_i weights correspond to the expenditure share on good *i* near the point of approximation.

To understand Rogoff's PPP puzzle, classify the *N* goods into the *K* traded goods, and $N - K$ nontraded goods, and then note we can write the log real exchange rate, q_t, as:

$$(2) \qquad q_t \equiv p_t^* - p_t = x_t + y_t ,$$

where x_t is the relative traded price indexes between the home and foreign country:

$$x_t \equiv p_{Tt}^* - p_{Tt} , \text{ with } \; p_{Tt} \equiv \sum_{i=1}^{K} \frac{\alpha_i}{\alpha_T} p_{it} , \; p_{Tt}^* \equiv \sum_{i=1}^{K} \frac{\alpha_i}{\alpha_T} p_{it}^* , \; \alpha_T \equiv \sum_{i=1}^{K} \alpha_i ,$$

and, where y_t involves the relative price of nontraded to traded goods at home and abroad:

$$y_t \equiv (1-\alpha_T)(p^*_{Nt} - p^*_{Tt} - (p_{Nt} - p_{Tt})), \text{ with } \ p_{Nt} \equiv \sum_{i=K+1}^{N} \frac{\alpha_i}{1-\alpha_T} p_{it}, \ p^*_{Nt} \equiv \sum_{i=K+1}^{N} \frac{\alpha_i}{1-\alpha_T} p^*_{it}.$$

Real tastes and technology shocks drive the relative price, y_t. Since those shocks are likely to be very persistent, the slow adjustment of real exchange rates could be explained if the y_t were predominant in movements in q_t. But as Rogoff notes, the innovation variance of y_t is much smaller than the innovation variance of x_t. This relative price component cannot account for the volatility in real exchange rates.

The x_t component of course is highly volatile. Its volatility is easy to understand in the context of models in which nominal consumer prices are sticky (when expressed in their own currencies.) In that case, x_t inherits the volatility of nominal exchange rates. As Rogoff notes, nominal exchange rate volatility can be explained by financial factors. But how do we explain the persistence of x_t? We might hypothesize that nominal prices adjust gradually toward their long run mean, perhaps according to an AR(1) process:

$$x_{it} = \theta x_{it-1} + \varepsilon_{it}, \text{ where } x_{it} \equiv p^*_{it} - p_{it}, \text{ and } \varepsilon_{it} \text{ is} \sim \text{i.i.d., } E(\varepsilon_{it}) = 0, \ E(\varepsilon^2_{it}) = \sigma^2.$$

Under these assumptions, x_t itself is an AR(1) process, with first-order serial correlation given by θ. But Rogoff notes that the persistence of real exchange rates (a half life of 3-5 years) is inconsistent with plausible estimates of the speed of nominal price adjustment.

Imbs et. al. (2002) (hereinafter IMRR) offer a potential resolution to this puzzle. They argue that the real exchange rate behavior can be reconciled with sticky-price models. Their explanation involves introducing heterogeneity into the processes for x_{it}:

(3) $$x_{it} = \theta_i x_{it-1} + \varepsilon_{it}, \text{ where } \varepsilon_{it} \text{ is} \sim \text{i.i.d, } E(\varepsilon_{it}) = 0, \ E(\varepsilon^2_{it}) = \sigma^2_i.$$

They argue that there is an "aggregation bias" when we use the half-life of real exchange rates to draw inferences about the speed of price adjustment. When this bias is corrected, they claim that estimates of price adjustment drawn from real exchange rates are completely in line with models of slow nominal price adjustment. They find an average half-life of price adjustment on the order of one year.

Our objective is to reexamine the case for "aggregation bias". In section 2, we restate IMRR's definition of "aggregation bias", and examine the statistical and economic meaning of this bias. Section 3 performs some simulations to assess the size of this bias. Our simulations differ from those in IMRR in that we restrict x_{it} to be non-explosive – that is $\theta_i \leq 1$. In section 4, we discuss measurement error. In general, x_{it} may appear to be less persistent than it actually is if its measurement is marred by an additive error that is not very persistent. Unfortunately, the data in IMRR has many such problems. Finally, as is now well known, there can be severe small sample bias in estimating autocorrelation coefficients. IMRR do not employ a correction for this bias, but we show its importance in section 5.

All four sections point toward the same direction: that "aggregation bias" is not a likely solution to the PPP puzzle. Section 2 shows that there is no general sense in which aggregation biases downward estimates of the speed of adjustment. The simulations of section 3 show that in practice the effect of aggregation bias is very small. Section 4 finds that when the data is cleaned, there is no empirical evidence of aggregation bias. Finally, the estimates of section 5 that correct for small sample bias imply very slow adjustment of prices – i.e., the PPP puzzle is worse than you think.

IMRR also investigate biases associated with another sort of heterogeneity. Panel studies of PPP frequently impose the identical speed of adjustment across all real exchange rates, but IMRR argue that that assumption can bias estimates of the half-life of real exchange rates. We do not address biases in panel estimates here – there is already a large literature devoted to that topic. While our empirical work replicates the panel methods of IMRR, our focus is on the aggregation bias.

2. Heterogeneity and Aggregation Bias

IMRR investigate the properties of q_t. In their empirical work, and their simulations, they make the following claim: the estimated half-life of q_t is much larger than the average estimated half-life of the x_{it}, where $q_t = \sum_{i=1}^{N} \alpha_i x_{it}$ and $\sum_{i=1}^{N} \alpha_i = 1$. Analytically, they demonstrate a different result: the first autocorrelation of q_t is larger than the average of the first autocorrelations of the x_{it}, when the x_{it} follow AR(1) processes defined in equation (3). First, we note that there is not necessarily a direct relationship between the size of the first autocorrelation and the half-life. An average of *N* variables that each are AR(1) generally will be an ARMA(*N*,*N*-1) process. The impulse response function for an ARMA(*N*,*N*-1) process can be non-monotonic and is not a simple function of the first autocorrelation.

Nonetheless, it is revealing to examine the expression for the first autocorrelation of q_t. It is not true that in general there is aggregation bias. That is, the first autocorrelation may be greater or less than the average of the first autocorrelations of the x_{it}. We can unambiguously state there is aggregation bias only when the α_i (the weights in the price index) are equal for all *i*; σ_i^2 (the innovation variance for x_{it} as defined in equation (3)) is the same for all *i*; and, the cross-correlations of all series are equal ($\sigma_{ij} \equiv E(\varepsilon_i \varepsilon_j)$, $i \neq j$, are all equal.)

Granger (1980) examines the properties of sums of AR series. Here we make some observations relevant for the discussion in IMRR.

The expression for the "aggregation bias" in the first autocorrelation is a nonlinear function of the autocorrelation coefficients θ_i, the weights α_i, the variances σ_i^2, and the covariances σ_{ij}.

Define $\theta^a \equiv \frac{Cov(q_t, q_{t-1})}{Var(q_t)}$, which is the first autocorrelation of the real exchange rate. The weighted average of the autocorrelation coefficients for the x_{it} is given by

$$\bar{\theta} \equiv \sum_{i=1}^{N} \alpha_i \theta_i .$$

As shown by Imbs, Mumtaz, Ravn and Rey (2003), the first order autocorrelation of the real exchange rate is given by:[1]

$$\theta^a = \bar{\theta} + \sum_{i=1}^{N} \lambda_i (\theta_i - \bar{\theta}), \tag{4}$$

$$\lambda_i = \left[\frac{\alpha_i^2 \sigma_i^2}{1-\theta_i^2} + \sum_{j=1, j\neq i}^{N} \frac{\alpha_i \alpha_j \sigma_{ij}}{1-\theta_i \theta_j} \right] \Bigg/ \left[\sum_{i=1}^{N} \frac{\alpha_i^2 \sigma_i^2}{1-\theta_i^2} + 2 \sum_{i=1}^{N} \sum_{j>i}^{N} \frac{\alpha_i \alpha_j \sigma_{ij}}{1-\theta_i \theta_j} \right].$$

"Aggregation bias " is defined as $\theta^a - \bar{\theta}$. Aggregation bias may be positive or negative. There is upward aggregation bias, as IMRR assert, if the higher values of θ_i get higher weights, λ_i, so that $\sum_{i=1}^{N} \lambda_i (\theta_i - \bar{\theta})$ is positive.

The first issue to address is the fact that the consumer price indexes are not simple unweighted averages of prices. The IMRR methodology imposes the assumption that the price

[1] This equation is essentially the same as the one derived in Imbs et. al. (2003) for equal weights. We have followed their nice derivation in Appendix A. As Imbs et. al. (2003) note, their version of equation (4) for *N* prices was derived in response to correspondence from one of the authors of this paper, which outlined the effects of non-zero correlation and different variances on aggregation bias that we describe here.

index is equally weighted.[2] Even if the aggregation bias (as defined here) is zero, θ^a does not necessarily equal the *unweighted* average of the θ_i.

This alone could lead IMRR to overstate aggregation bias. Rogers and Jenkins (1995) found that they were only able to reject the null that x_{it} has a unit root for specific fresh food items – eggs, oranges, apples, bananas, chuck roast, and a few others. But fresh food has a low weight in the overall CPI. Simply taking an unweighted average of speeds of adjustment gives too much weight to these goods, and makes the unweighted average of the θ_i lower than the weighted average, $\bar{\theta}$.

In the special case in which the weights are all equal ($\alpha_i = N^{-1}$), the innovation variances for all x_{it} are equal ($\sigma_i^2 = \sigma^2$), and all of the covariances of innovations are equal ($\sigma_{ij} = \rho\sigma^2$), Imbs et. al. (2003) show that aggregation bias is positive.

In general, however, aggregation bias could be positive or negative. We can write:

$$\lambda_i = \frac{Cov(\alpha_i x_{it}, q_t)}{Var(q_t)} = \frac{Cov(\alpha_i x_{it}, \sum_{j=1}^{N} \alpha_j x_{jt})}{Var(\sum_{j=1}^{N} \alpha_j x_{jt})}.$$

For example, if the series were uncorrelated and weights were equal, $\lambda_i = \dfrac{Var(x_{it})}{\sum_{j=1}^{N} Var(x_{jt})}$. If the variance of x_{it} is large relative to the average variance, then θ_i will be disproportionately weighted in θ^a. But that could cause θ^a to be too large (if θ_i is larger than $\bar{\theta}$) or too small (if θ_i is smaller than $\bar{\theta}$).

[2] For example, their RCM estimator (discussed below) has the mean speed of adjustment as an unweighted average of the θ_i.

More generally, if the series are correlated but the weights are equal, $\lambda_i = \dfrac{Cov(x_{it}, q_t)}{N \cdot Var(q_t)}$.

Allowing for non-zero correlation could push the aggregation bias in any direction. Consider this simple example. Suppose $N = 3$, weights are equal, and there is perfect negative correlation between x_{1t} and x_{2t}: specifically, $x_{1t} = -x_{2t}$. Then $q_t = \frac{1}{3} x_{3t}$. The persistence of the real exchange rate is the same as the persistence of x_{3t}.

The aggregation bias is a nonlinear function of the parameter vectors $\vec{\theta}$, $\vec{\alpha}$, and the variance-covariance matrix, Σ. That is, $\theta^a = f(\vec{\theta}, \vec{\alpha}, \Sigma)$. Take a first-order Taylor series approximation of $f(\vec{\theta}, \vec{\alpha}, \Sigma)$, around a point of $\theta_i = \theta$ for all *i*. It follows from (4) that

$$\theta^a = \theta + \sum_{i=1}^{N} \bar{\lambda}_i (\theta_i - \theta),$$

where $\bar{\lambda}_i$ is the value of λ_i evaluated at the point of expansion. There are two special cases to consider. First, suppose we evaluate $f(\vec{\theta}, \vec{\alpha}, \Sigma)$ at a point where all weights are equal, all variances are equal, and all covariances are equal. We find (see Appendix A):

$$\theta^a = \frac{1}{N} \sum_{i=1}^{N} \theta_i .$$

That is, near the point of expansion, there is no aggregation bias. Alternatively, if we evaluate at a point where the weights are not equal, but the variances of the innovations are equal and they are perfectly correlated, we find

$$\theta^a = \sum_{i=1}^{N} \alpha_i \theta_i = \bar{\theta} .$$

Again, there is no aggregation bias. In the data, the innovations in the x_{it} are all highly correlated with the innovations in the nominal exchange rate.

In other words, aggregation bias is second order. Unless there are very large differences in speeds of adjustment, large differences in innovation variances, big differences in weights, and low correlation, aggregation bias will not be large.

In the next section, we consider simulations to gauge the size of the aggregation bias.

3. Simulations

We first examine aggregation bias in simple OLS simulations. An example of the type of simulation we perform is: generate *N* AR(1) series with sample size *T* = 50, 100, 150, 200, 250, 300 and 500. The AR(1) coefficients in these series are drawn from a uniform distribution on the range of 0.881 to 0.999 with mean $\theta = 0.94$. We note that if a series has an AR(1) coefficient of 0.94, its half life is 11.2 periods. We choose $\theta = 0.94$ to match the baseline simulations in IMRR.

More precisely, the data generating processes is:

$$x_{it} = \varsigma_i + \theta_i x_{it-1} + \varepsilon_{it},$$

with two different specifications:

1. **DGP1:** $\varsigma_i = 0$, $\theta_i = \theta + \mu_i$, $\theta = 0.94$, $\mu_i \sim^{i.i.d.} uniform[-0.059, 0.059]$,
2. **DGP2:** $\varsigma_i \sim^{i.i.d.} N(0,1)$, $\theta_i = \theta + \mu_i$, $\theta = 0.94$, $\mu_i \sim^{i.i.d.} uniform[-0.059, 0.059]$.

Then average the *N* series and estimate an AR(1) on the average series. Let θ^a be the autoregressive coefficient for the average of the x_{it}. To calibrate the aggregation bias per se, we average the estimates of the θ_i from the *N* series and compare it to the estimated θ^a, which we call $\hat{\theta}^a$. Each experiment is conducted with 2500 draws.

These experiments are designed to meet the criterion of the theorem of Imbs et. al. (2003) that guarantees that aggregation bias will make the autoregressive coefficient of the average series greater than the average of the autoregressive coefficients from the individual series. That is, each series is weighted equally in the average, the innovation variances are equal, and the cross-correlations are all equal (set equal to zero.)

Table 1 reports results with $N = 100$. We first direct attention to the column labeled "Bias Uncorrected". Perhaps surprisingly, the estimated AR(1) coefficient on the average price is usually less than the mean θ from the individual series. There does not appear to be aggregation bias unless we allow for a very long sample period. Otherwise, the bias goes the other direction.

The reason for the downward bias is the familiar problem of small-sample bias. The small sample bias trumps the aggregation bias unless the series is very long. (This matches the finding of Choi, Mark, and Sul (2003) with similar simulations.)

Next, we investigate if the number of individual price series affects our results. We thus increase the number of AR(1) series from 100 to 500. Clearly, Table 2 indicates that the results are similar even when we aggregate more individual series.

We then consider a new DGP with higher persistence (θ=0.98) and present the results in Table 3.[3] The estimated coefficient on the average series is almost always less than 0.98.

In Table 4, we examine how the degree of cross-sectional heterogeneity in persistence affects the magnitude of aggregation bias. We set the number of sectors equal to 100, sample size equal to 300, and θ=0.94. μ_i is assumed to be distributed uniformly in the ranges of [-0.009,0.009], [-0.019,0.019], [-0.039,0.039] and the benchmark case: [-0.059,0.059]. Aggregation bias is larger when the degree of heterogeneity is higher.

In each table, we also attempt to separate out the effects of small-sample bias from aggregation bias. There is no aggregation bias if $\theta_i = \theta$ for all *i*. So for each table, we redo all

[3] We need to assume μ_i ~uniform[-0.019,0.019] to keep θ_i <1.

simulations imposing $\theta_i = \theta$ for all *i*. Call this estimate $\hat{\theta}$. That gives us an estimate of the small sample bias as simply $\hat{\theta} - \theta$. We then subtract the small sample bias from our estimate of $\hat{\theta}^a$ (the autoregressive coefficient of the average series) to get our Bias Corrected estimate, $\hat{\theta}^a_{bc} = \hat{\theta}^a - (\hat{\theta} - \theta)$. We report these bias corrected estimates of the autoregressive coefficient in columns 5-8 of each table. Comparing these to the true θ used to generate the data gives us a measure of the aggregation bias that is unpolluted by small-sample bias.

We find that there is some aggregation bias, but it is not nearly so dramatic as the simulations of IMRR indicate. For example, consider our base simulations in Table 1. With $\theta = 0.94$, (implying a half-life of 11.2 months), we find that $\hat{\theta}^a_{bc} \cong 0.956$ when the sample size is 250. This implies a half-life of 15.4 months. In the empirical work of IMRR, the sample sizes closer to 150 months for most of the series. Then we find $\hat{\theta}^a_{bc} \cong 0.952$, implying a half-life of 14.1 months. None of the simulations show a very large aggregation bias.

Aggregation bias appears not to be a major problem in this case at least in part because we know that the first-order serial correlation of the aggregated series is so large in the data. Since the autocorrelations of the individual series are bounded above by one, we cannot observe a great deal of heterogeneity among these correlations and still generate an aggregate with a serial correlation of 0.94. The aggregation bias in this case is not large because there is not much heterogeneity to be aggregated.

Why do we find that aggregation bias is so small in our simulations (and almost always dominated by small-sample bias) while the simulations in IMRR indicate a strikingly larger aggregation bias? In generating the AR(1) random variables with a distribution of first-order autoregressive coefficients, the simulations in the IMRR fail to put an upper bound of one on that coefficient. That is, they use series that are explosive, which is an inappropriate representation for the log of real exchange rates. (Granger (1980) states explosive roots are "generally considered to be inappropriate for economic variables.") The problem this causes is clear. For example, if there is one series with a root of 1.08 (which is the upper bound in some of the simulations in IMRR), and that runs for 100 observations, it will explode to such a degree that when it is averaged in with the stationary and unit root series, it will completely dominate the average. (1.08 to the 100^{th} power equals around 2200.) The average will look a lot like an AR(1) with a root of 1.08.

4. Data and Measurement Error

In order to investigate the importance of parameter heterogeneity, we use relative prices of goods at the sectoral level for 16 categories of goods for the United States and 9 European countries: Belgium, Denmark, Germany, Spain, France, Italy, Netherlands, Portugal, and the United Kingdom. Our monthly price data is obtained from Eurostat and is the data that was used in Engel (2000), although we have updated the cleaned-up data to 1996:12 to match the IMRR data set. The length of each series varies. The longest runs from 1981:1 to 1996:12 while the shortest runs from 1981:1 to 1994:9. The nominal exchange rate data is collected from the International Financial Statistics (IFS) database. A comprehensive description of data is provided in Appendix C.

Comparing our data set with what was used in IMRR, there are some issues worth noting: First, an odd thing about the IMRR data set is that it is missing data on nominal exchange rates from 1976-1991 or 1993 for six countries: Denmark, Finland, Greece, Ireland, Portugal, and the Netherlands. Simply by using a widely available source for exchange rates – the IFS database – their data sample can be considerably enlarged. The IFS data on exchange rates is virtually identical to the data IMRR use for those data points that are in both data sets. Second, our cleaned up price data only starts in 1981:1, while the IMRR uncorrected data starts in 1975:1. However, that data has few observations before 1981:1; that is, there are mostly missing observations from 1975:1-1980:12 in their uncorrected data.

Figures 1 and 2 in Appendix D compare a few of the series used by IMRR with our cleaned data. We follow Engel (2000) in cleaning the data. As the figures indicate, in many of the series there are a very small number of observations that apparently have very large errors. Almost all the data we use, and all the corrections, coincide with the data in Engel (2000). That paper (fn. 1, p. 1453) describes the data-cleaning process: "In cases where there were more than a few errors in one series, the series was dropped. Hence, there is not data for all nine countries for all of the goods. In cases where data entry errors were near the beginning or end of the series, the series were truncated. In some cases, the data was corrected from other sources. In some cases the entry error involved a transposition of digits which was corrected. In the remaining cases, the data points were replaced using interpolation from adjacent data points. The total number of data points used in the tables that are corrected data is 99 (out of a total of approximately 40,000 data points.)"

Measurement error will reduce the measured persistence of a stationary series if the measurement error is less persistent than the series itself, assuming the measurement error is uncorrelated with the true series. That is, let $y_t = q_t + z_t$, where q_t is the series whose persistence we would like to measure, and z_t is measurement error. If we measure persistence by the first autocorrelation, then

$$\frac{\mathrm{cov}(y_t, y_{t-1})}{\mathrm{var}(y_t)} = \frac{\mathrm{cov}(q_t, q_{t-1}) + \mathrm{cov}(z_t, z_{t-1})}{\mathrm{var}(q_t) + \mathrm{var}(z_t)},$$

assuming independence of q_t and z_t. But then, if $\frac{\operatorname{cov}(q_t, q_{t-1})}{\operatorname{var}(q_t)} > \frac{\operatorname{cov}(z_t, z_{t-1})}{\operatorname{var}(z_t)}$, we conclude that the measured series is less persistent than the true series: $\frac{\operatorname{cov}(y_t, y_{t-1})}{\operatorname{var}(y_t)} < \frac{\operatorname{cov}(q_t, q_{t-1})}{\operatorname{var}(q_t)}$. On the other hand, the average price series will be plagued less by the bias introduced by measurement error, if the measurement error is independent across prices (but innovations in prices are correlated.)

The data from IMRR plotted in Figures 1 and 2 (more plots are available from the authors on request) appear to have very large, very transitory measurement error. In some cases, the corrections made by Engel (2000) may smooth the data too much (when the bad data were corrected by interpolation.) But the number of data points smoothed in this way – indeed, the number of foul data points altogether – is small. It is unlikely that the cleaning process has much of an effect in making the data appear more persistent than it is. But clearly the uncleaned data could make the series appear much less persistent – even a small number of errors of the size evident in Figure 1 and 2 can have a big impact on the estimates of the autoregressive coefficients.

5. Empirical Estimates

5.a. Correcting small-sample bias

Small sample bias will lead to underestimation of half-life of deviations from PPP. Murray and Papell (2002a, 2002b, and 2003) address the problem of small sample bias in both univariate time series and panel data. Applying the median-unbiased estimation methods proposed by Andrews (1993) and Andrews and Chen (1994), they demonstrate that most recent studies which report shorter half-lives than the consensus in long-horizon data or panel data underestimate the half-lives of PPP deviations, and thus overestimate the speed of mean reversion. Most of the unbiased point estimates of half-lives lie within the 3-5 year range and in some cases, even longer. They conclude that "panels do not help solve the purchasing power parity puzzle," and "the purchasing power parity puzzle is worse than you think".

Cashin and McDermott (2003) also calculate the median-unbiased half-lives of parity deviations in three different models, Dickey-Fuller regression, Augmented Dickey-Fuller regression, and Phillips-Perron regression to investigate the real effective exchange rate for 22 industrial countries. Their findings come to the same conclusions as in Murray and Papell (2002a, 2002b): after removing the downward bias of standard autoregressive estimators, the cross-country average of half-lives of deviations from parity range between 4-15 years.

Many different approaches have been suggested to correct for small sample bias. The first one is the median-unbiased estimator proposed by Andrews (1993) and Andrews and Chen (1994). Andrews (1993) shows how to obtain the exactly median-unbiased estimates and exact confidence intervals of half-lives in a AR(1) model. For autoregressive model with higher order, Andrews and Chen (1994) demonstrate how to perform approximately median-unbiased estimation of AR coefficients. The efficiency of this method relies on the availability of the known median function and precise distribution assumptions. Moreover, median function does not work well if the true AR(1) coefficient is near unity, which is the case when studying PPP since real exchange rates exhibit very high persistence.

A second approach is mean-unbiased estimates, proposed by Kilian (1998). This method uses a sample reuse procedure based on a bootstrap. The bias-corrected AR coefficients and confidence intervals for impulse response function are obtained by using a bootstrap-after-bootstrap method. Since Kilian's bias-corrected percentile method is designed and motivated for stationary autoregressions, it may perform worse when the process is near nonstationary.

A third approach relies on recursive demeaning procedures, which is introduced by So and Shin (1999) and has be applied in Sul, Phillips and Choi (2003) and Choi, Mark and Sul (2003). The idea of this method is that the recursive demeaning reduces the bias from the correlation between residuals and the lagged dependent variables when fitting an intercept. The advantage of this method over the other two methods mentioned above is that it is simple to apply. The correction procedure only requires the sample mean to be replaced by the partial sample mean (the average of data up to the time point) when conducting estimations. Since the approximately median-unbiased estimators proposed by Andrews and Chen (1994) are computationally intense, it is less feasible for our purpose to correct the small sample bias in a large panel data set with number of group (N) being about 130. We will apply the second and third bias correction methods (i.e. the Kilian (1998)'s method and So and Shin (1999)'s method) to correct the small sample bias in the group mean estimations. The details of group mean estimators (random coefficients model and mean group estimator) and the bias correcting procedures are presented in Appendix B.

5.b. Empirical Results

In this section, we present our empirical results of the random-coefficient (RCM) and mean-group (MG) estimations under bias corrections. In order to compare with the estimates in IMRR, we also conduct the estimations without bias correction. The half-lives are obtained from impulse response functions and the 95% confidence intervals for the half-lives are calculated using nonparametric bootstrap with 1000 replications. The orders of autoregression are chosen to be 5 to match those repored in IMRR.[4] Our estimation results are reported in Table 5.

First, consider column (1), which contains the sum of AR coefficients, half-life, and 95% confidence interval for the half-life in the bias-uncorrected RCM and MG estimations. The bias-

[4] Our persistence results are nearly identical when we use 4, 5, or 6 lags.

uncorrected estimates of MG estimator suggest a range of the half-lives around 25-26 months, which is below Rogoff's 36 to 60 months consensus. The half-life estimates from RCM model, however, are very close to the consensus.

We note that we are able to replicate IMRR empirical results using their uncorrected data set quite closely. Our point estimates of the sums of the autoregressive parameters are very similar. In the RCM model, our point estimate is 0.9311, which compares to a point estimate of 0.9481 reported in IMRR for the RCM estimator. We find a half-life of 12 months compared to IMMR's 14 months. So, though our dataset has dropped a few of the prices used in IMRR, when we use the uncorrected data, we find the very low half-lives that they do. (Indeed our data exhibit slightly shorter half-lives.) But these short half-lives disappear entirely when we use the corrected data.[5]

As mentioned above, these GLS weighted/simple weighted least squares estimators of AR coefficients are underestimated because of downward small sample bias. We now report the bias-corrected estimates using the correcting procedure suggested by Kilian (1998) and So and Shin (1999) in columns (2) and (3). The results in column (2) or (3) suggest that correcting the small sample bias raise the point estimates of the half-lives in both RCM and MG estimations. They are well above 3 years. When So and Shin's (1999) method is used, the half-life estimates of PPP deviations are about 17 years for RCM and are about 13 years for MG. These results are consistent with a much slower rate of convergence to PPP as found in Murray and Papell (2002b) and Choi, Mark and Sul (2003).

5.c. Country-by-country study

[5] Our longer half-lives are coming both because we correct the data errors, and because we supply the missing exchange rate data. With our exchange rates, but the uncorrected price data, we find the sum of the AR coefficients to be .9662, and a half-life of 21 months.

It may not be appropriate to do panel estimates across countries if the main concern is the magnitude of aggregation bias. Aggregation bias means that aggregating the goods into aggregated price indices causes a bias in estimates of the speed of adjustment of the real exchange rate when the individual goods adjust at different speeds. However, the other well-known problem in panel estimations is that estimating in a panel a single mean speed of adjustment for different countries may cause bias as well. A panel study across countries mixes the two problems, which means that we really cannot tell how much of the problem comes from the price indexes aggregation versus the problem arising from imposing the same mean speed of adjustment for each country. To concentrate on the aggregation problem, we investigate the country-by-country estimates of the bias-corrected speed of adjustment of real exchange rates in both aggregated and disaggregated level. We can then measure the aggregation bias more precisely than a cross-country panel study. The results are presented in Table 6. Here we report the results after correcting small sample bias by So and Shin (1999)'s method with 5 lags. Using Kilian (1998)'s method and different lag lengths gives us similar results.

The first two columns are results using aggregated data. Column 1 (AGG-CPI) simply uses consumer price index (CPI) data while column 2 (AGG-TCPI) contains the results from equally weighted aggregation of our cross-sectional price data. There is only a slight difference between the point estimates of persistence from different measures of aggregation. Columns 3 and 4 use disaggregated price data with MG estimation (column 3) and RCM estimation (column 4). Clearly, by comparing columns 1 and 2 with columns 3 and 4, it can be found that aggregation bias is small in most countries. Furthermore, in the case of the RCM estimator, there is no aggregation bias, no matter which method of aggregation is used (AGG-CPI vs. AGG-TCPI).

6. Concluding Remarks

According to Rogoff (1996), the consensus of empirical studies is that half-life deviations from purchasing power parity are between 3 and 5 years. The documented low convergence speed in real exchange rates induces a purchasing power puzzle since the half-lives of PPP deviations seem too long to be justified by models with nominal rigidities.

IMRR (2002) demonstrate the importance of the role of heterogeneity and aggregation bias in explaining the PPP puzzle. Based on 2-digit sub-sectional price data from 19 categories of goods and 13 countries, they find much faster mean reversion of the real exchange rate than the consensus view. They claim that considering aggregation bias solves the PPP puzzle.

In this paper, we have re-examined the claim in IMRR. We investigate the same data set in IMRR with corrections of a few entry errors in the data and additions of the missing nominal exchange rate data in their study. Using the cleaned data set, we find half-life estimates in line with Rogoff's consensus. Further, using two different bias-correcting methods proposed by Kilian (1998) and So and Shin (1999), we find that the point estimates of half-life deviations from PPP are even higher than Rogoff's consensus. We cast doubt on the claim that aggregation bias can account for the PPP puzzle.

Moreover, we argue based on theory and simulations, that aggregation bias is unlikely to be large.

The finding that aggregation bias is not large should be no surprise to readers of the literature. Engel (2000) uses the same data source as IMRR (but with cleaned data) and concludes that for most goods adjustment is slow. That paper concludes, "for most categories of goods, there is not even evidence that deviations from the law of one price tend to be eliminated." This is also similar to the conclusion reached by Rogers and Jenkins (1995), examining the behavior of relative prices between the U.S. and Canada for a long list of narrowly defined goods. In very few cases could they even reject the null of a unit root in the deviations from the law of one price. Crucini and Shintani (2002), using a large worldwide cross-section of goods prices from the Economics Intelligence Unit, do find rapid convergence to deviations from the law of one price (half-lives of 9 to 12 months.) But they explicitly find no aggregation bias. The rate of convergence of the average of their prices is very similar to the average of the rates of convergence of their individual prices.

Aggregation bias does not seem to explain the PPP puzzle.

For Tables 1-4: The "Bias Corrected" for AGG (columns 5 and 7) is computed as (1) column 1 + small sample bias and (2) column 3 + small sample bias. The small sample bias is obtained as follows (1) generating series from the DGP with all of the θs are the same (no aggregation bias); (2) aggregating these series; (3) estimating the aggregated series; (4) small sample bias is the difference between the estimate of θ and the true θ.

Table 1: Simulation Results (Number of individual price series = 100, θ=0.94)

	Bias Uncorrected				Bias Corrected			
	DGP1		DGP2		DGP1		DGP2	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Sample Size	AGG	TRUE	AGG	TRUE	AGG	TRUE	AGG	TRUE
50	0.8574	0.9398	0.8573	0.9400	0.9457	0.9398	0.9458	0.9400
100	0.9065	0.9399	0.9071	0.9401	0.9483	0.9399	0.9502	0.9401
150	0.9240	0.9401	0.9248	0.9401	0.9520	0.9401	0.9527	0.9401
200	0.9342	0.9400	0.9339	0.9400	0.9541	0.9400	0.9555	0.9400
250	0.9402	0.9400	0.9410	0.9399	0.9562	0.9400	0.9565	0.9399
300	0.9445	0.9400	0.9446	0.9401	0.9581	0.9400	0.9585	0.9401
500	0.9541	0.9399	0.9540	0.9399	0.9615	0.9399	0.9613	0.9399

Note: AGG is the average AR(1) coefficient estimated from average price. TRUE is the true average AR(1) coefficient.

Table 2: Simulation Results (Number of individual price series = 500, θ=0.94)

	Bias Uncorrected				Bias Corrected			
	DGP1		DGP2		DGP1		DGP2	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Sample Size	AGG	TRUE	AGG	TRUE	AGG	TRUE	AGG	TRUE
50	0.8537	0.9400	0.8566	0.9400	0.9434	0.9400	0.9432	0.9400
100	0.9078	0.9400	0.9074	0.9399	0.9502	0.9400	0.9501	0.9399
150	0.9246	0.9400	0.9251	0.9400	0.9524	0.9400	0.9527	0.9400
200	0.9361	0.9400	0.9351	0.9401	0.9554	0.9400	0.9553	0.9401
250	0.9410	0.9400	0.9412	0.9400	0.9574	0.9400	0.9570	0.9400
300	0.9445	0.9400	0.9448	0.9400	0.9582	0.9400	0.9578	0.9400
500	0.9542	0.9400	0.9540	0.9400	0.9619	0.9400	0.9615	0.9400

Note: AGG is the average AR(1) coefficient estimated from average price. TRUE is the true average AR(1) coefficient.

Table 3: Simulation Results (Number of individual price series = 100, θ=0.98)

	Bias Uncorrected				Bias Corrected			
	DGP1		DGP2		DGP1		DGP2	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Sample Size	AGG	TRUE	AGG	TRUE	AGG	TRUE	AGG	TRUE
50	0.8858	0.9800	0.8813	0.9800	0.9802	0.9800	0.9806	0.9800
100	0.9341	0.9800	0.9345	0.9800	0.9806	0.9800	0.9810	0.9800
150	0.9492	0.9800	0.9490	0.9800	0.9810	0.9800	0.9813	0.9800
200	0.9590	0.9800	0.9584	0.9800	0.9818	0.9800	0.9817	0.9800
250	0.9647	0.9800	0.9643	0.9800	0.9826	0.9800	0.9820	0.9800
300	0.9678	0.9800	0.9681	0.9800	0.9834	0.9800	0.9834	0.9800
500	0.9752	0.9800	0.9757	0.9800	0.9842	0.9800	0.9842	0.9800

Note: AGG is the average AR(1) coefficient estimated from average price. TRUE is the true average AR(1) coefficient.

Table 4: Simulation Results (Number of individual price series = 100, θ=0.94, Sample Size = 300)

	Bias Uncorrected				Bias Corrected			
	DGP1		DGP2		DGP1		DGP2	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
μ_i~uniform[-a,a]	AGG	TRUE	AGG	TRUE	AGG	TRUE	AGG	TRUE
[-0.009,0.009]	0.9268	0.9400	0.9275	0.9400	0.9403	0.9400	0.9403	0.9400
[-0.019,0.019]	0.9281	0.9400	0.9284	0.9400	0.9415	0.9400	0.9417	0.9400
[-0.039,0.039]	0.9334	0.9401	0.9332	0.9401	0.9470	0.9401	0.9468	0.9401
[-0.059,0.059]	0.9439	0.9399	0.9442	0.9400	0.9578	0.9399	0.9577	0.9400

Note: AGG is the average AR(1) coefficient estimated from average price. TRUE is the true average AR(1) coefficient.

Table 5: Half-life Estimations. Monthly data from 1981:1 to 1996:12.

$$x_{it} = \varsigma_i + \sum_{k=1}^{K} \theta_{ik} x_{it-k} + \nu_{it} \text{ with } \theta_{ik} = \theta_k + \xi_{1i} \text{ and } \varsigma_i = \varsigma + \xi_{2i}.$$

	(1) Bias-Uncorrected			(2) Bias-Corrected, Kilian (1998)			(3) Bias-Corrected, So and Shin (1999)		
	$\sum_{k=1}^{K}\theta_k$	HL	95%CI	$\sum_{k=1}^{K}\theta_k$	HL	95%CI	$\sum_{k=1}^{K}\theta_k$	HL	95%CI
RCM	0.9783	34	[20,118]	0.9832	43	[12,∞]	0.9966	215	[111, ∞]
MG	0.9706	26	[20,142]	0.9836	44	[13,∞]	0.9955	161	[112, ∞]

Table 6: Country-by-Country Half-life Estimations. Monthly data from 1981:1 to 1996:12.

	CPI		TCPI		MG		RCM	
	$\sum_{k=1}^{K}\theta_k$	HL [95%CI]	$\sum_{k=1}^{K}\theta_k$	HL [95%CI]	$\sum_{k=1}^{K}\theta_k$	HL [95%CI]	$\sum_{k=1}^{K}\theta_k$	HL [95%CI]
Belgium	0.9825	42	0.9804	38	0.9702	26	1.0190	∞
		[32,63]		[31,56]		[24,31]		[13, ∞]
Denmark	0.9859	52	0.9805	38	0.9766	31	0.9900	73
		[35,96]		[28,61]		[23, ∞]		[12, ∞]
Germany	0.9820	41	0.9729	28	0.9686	24	0.9734	28
		[19,120]		[17,49]		[20,55]		[17, ∞]
Spain	0.9886	63	0.9843	46	0.9852	47	0.9904	74
		[47,100]		[36,67]		[30, ∞]		[9, ∞]
France	0.9791	36	0.9737	29	0.9618	20	0.9756	31
		[26,57]		[22,46]		[18,27]		[13, ∞]
Italy	0.9821	42	0.9758	31	0.9693	25	0.9974	283
		[34,56]		[25,41]		[20,58]		[8, ∞]
Netherlands	0.9770	32	0.9601	19	0.9730	28	0.9850	49
		[18,77]		[14,32]		[21,780]		[17, ∞]
Portugal	0.9909	81	0.9903	76	0.9841	44	1.0173	∞
		[62,140]		[58,130]		[39,71]		[13, ∞]
UK	0.9660	23	0.9306	13	0.9614	20	0.9653	22
		[14,40]		[9,19]		[18,38]		[16, ∞]

References

Andrews, Donald W K (1993). "Exactly Median-Unbiased Estimation of First Order Autoregressive /Unit Root Models," *Econometrica*, vol. 61, no. 1, pp. 139-65.

Andrews, Donald W K and Hong-Yuan Chen (1994). "Approximately Median-Unbiased Estimation of Autoregressive Models," *Journal of Business and Economic Statistics*, vol. 12, no. 2, pp. 187-204.

Cashin, Paul A. and C. John McDermott (2003). "An Unbiased Appraisal of Purchasing Power Parity," *IMF Staff Paper*, forthcoming.

Choi, Chi-Young, Nelson C. Mark and Donggyu Sul (2003). "The Dominance of Downward Bias in Half-Life Estimates of PPP Deviations," working paper, University of Notre Dame.

Crucini, Mario J., and Mototsugu Shintani (2002). "Persistence in Law-of-One-Price Deviations: Evidence from Micro-data," working paper, Vanderbilt University.

Engel, Charles (2000). "Local-Currency Pricing and the Choice of Exchange-Rate Regime," *European Economic Review*, vol. 44, no. 8, pp. 1449-72.

Granger, Clive W.J. (1980). "Long Memory Relationships and the Aggregation of Dynamic Models," *Journal of Econometrics*, vol. 14, no. 2, pp. 227-38.

Imbs, Jean, Haroon Mumtaz, Morten Ravn and Helene Rey (2002)."PPP Strikes Back: Aggregation and the Real Exchange Rate," NBER Working Paper 9372.

Imbs, Jean, Haroon Mumtaz, Morten Ravn and Helene Rey (2003). "The PPP Controversy - A Summary of the Debate Surrounding 'PPP Strikes Back'" www.london.edu/~jimbs

Kilian, Lutz (1998). "Small-Sample Confidence Intervals for Impulse Response Functions," *Review of Economics and Statistics*, vol. 80, no. 2, pp. 218-30.

Murray, Christian J and David H Papell (2002a). "The Purchasing Power Parity Persistence Paradigm," *Journal of International Economics*, vol. 56, no. 1, pp. 1-19.

Murray, Christian J and David H Papell (2002b). "The Purchasing Power Parity Puzzle is Worse Than You Think," working paper, University of Houston.

Murray, Christian J and David H Papell (2003). "Do Panels Help Solve the Purchasing Power Parity Puzzle?" working paper, University of Houston.

Rogers, John H., and Michael Jenkins (1995). "Haircuts or Hysteresis? Sources of Movements in Real Exchange Rates," *Journal of International Economics*, vol. 38, no. 3-4, pp. 339-60.

Rogoff, Kenneth (1996). “The Purchasing Power Parity Puzzle,” *Journal of Economic Literature*, vol. 34, no. 2, pp. 647-68.

So, Beong Soo and Dong Wan Shin (1999).”Recursive Mean Adjustment in Time-Series Inferences,” *Statistics and Probability Letters*, vol. 43, pp. 65-73.

Sul, Donggyu, P.C.B. Phillips and Chi-Young Choi (2002).“Prewhitening Bias in HAC Estimation,” working paper, The University of Auckland.

APPENDIX A

1. Derivation of Equation (4):[6]

Equation (A1) and (A2) represent our model:

(A1) $x_{it} = \theta_i x_{it-1} + \varepsilon_{it}$, $E(\varepsilon_{it}) = 0$, $E(\varepsilon_{it}^2) = \sigma_i^2$, $E(\varepsilon_{it}\varepsilon_{jt}) = \sigma_{ij}$,

(A2) $q_t = \sum_{i=1}^{N} \alpha_i x_{it}$, $\sum_{i=1}^{N} \alpha_i = 1$.

The least square estimate of the persistence of q_t is given by θ^a:

$$\theta^a = \frac{Cov(q_t, q_{t-1})}{Var(q_t)},$$

where

(A3) $$Cov(q_t, q_{t-1}) = \sum_{i=1}^{N} \left[\theta_i \alpha_i^2 \frac{\sigma_i^2}{1-\theta_i^2} + \sum_{j>i}^{N} (\theta_i + \theta_j) \alpha_i \alpha_j \frac{\sigma_{ij}}{1-\theta_i \theta_j} \right],$$

and

(A4) $$Var(q_t) = \sum_{i=1}^{N} \left[\alpha_i^2 \frac{\sigma_i^2}{1-\theta_i^2} + 2 \sum_{j>i}^{N} \alpha_i \alpha_j \frac{\sigma_{ij}}{1-\theta_i \theta_j} \right].$$

Hence, it follows that:

[6] The derivation follows that of Imbs. et. al. (2003), allowing for different weights in the price index.

$$\theta^a = \frac{\sum_{i=1}^{N}\left[\theta_i\alpha_i^2\frac{\sigma_i^2}{1-\theta_i^2}+\sum_{j>i}^{N}(\theta_i+\theta_j)\alpha_i\alpha_j\frac{\sigma_{ij}}{1-\theta_i\,\theta_j}\right]}{\sum_{i=1}^{N}\left[\alpha_i^2\frac{\sigma_i^2}{1-\theta_i^2}+2\sum_{j>i}^{N}\alpha_i\alpha_j\frac{\sigma_{ij}}{1-\theta_i\,\theta_j}\right]}$$

$$=\bar{\theta}-\bar{\theta}+\frac{\sum_{i=1}^{N}\left[\theta_i\alpha_i^2\frac{\sigma_i^2}{1-\theta_i^2}+\sum_{j>i}^{N}(\theta_i+\theta_j)\alpha_i\alpha_j\frac{\sigma_{ij}}{1-\theta_i\,\theta_j}\right]}{\sum_{i=1}^{N}\left[\alpha_i^2\frac{\sigma_i^2}{1-\theta_i^2}+2\sum_{j>i}^{N}\alpha_i\alpha_j\frac{\sigma_{ij}}{1-\theta_i\,\theta_j}\right]}$$

$$=\bar{\theta}+\frac{\sum_{i=1}^{N}\left[(\theta_i-\bar{\theta})\alpha_i^2\frac{\sigma_i^2}{1-\theta_i^2}+\sum_{j>i}^{N}[(\theta_i-\bar{\theta})\alpha_i\alpha_j\frac{\sigma_{ij}}{1-\theta_i\,\theta_j}+(\theta_j-\bar{\theta})\alpha_i\alpha_j\frac{\sigma_{ij}}{1-\theta_i\,\theta_j}]\right]}{\sum_{i=1}^{N}\left[\alpha_i^2\frac{\sigma_i^2}{1-\theta_i^2}+2\sum_{j>i}^{N}\alpha_i\alpha_j\frac{\sigma_{ij}}{1-\theta_i\,\theta_j}\right]}$$

$$=\bar{\theta}+\sum_{i=1}^{N}\left(\frac{\alpha_i^2\frac{\sigma_i^2}{1-\theta_i^2}+\sum_{j=1,\,j\neq i}^{N}\alpha_i\alpha_j\frac{\sigma_{ij}}{1-\theta_i\,\theta_j}}{\sum_{i=1}^{N}\left[\alpha_i^2\frac{\sigma_i^2}{1-\theta_i^2}+2\sum_{j>i}^{N}\alpha_i\alpha_j\frac{\sigma_{ij}}{1-\theta_i\,\theta_j}\right]}\right)(\theta_i-\bar{\theta}).$$

Define λ_i as

$$\lambda_i \equiv \frac{\frac{\alpha_i^2\sigma_i^2}{1-\theta_i^2}+\sum_{j=1,\,j\neq i}^{N}\frac{\alpha_i\alpha_j\sigma_{ij}}{1-\theta_i\theta_j}}{\sum_{i=1}^{N}\frac{\alpha_i^2\sigma_i^2}{1-\theta_i^2}+2\sum_{i=1}^{N}\sum_{j>i}^{N}\frac{\alpha_i\alpha_j\sigma_{ij}}{1-\theta_i\theta_j}},$$

we can get equation (4) in page 6:

$$\theta^a=\bar{\theta}+\sum_{i=1}^{N}\lambda_i(\theta_i-\bar{\theta}).$$

2. First-Order Taylor Series Approximation of θ^a

From equation (4),

$$\theta^a = f(\vec{\theta}, \vec{\alpha}, \Sigma) = \overline{\theta} + \sum_{i=1}^{N} \lambda_i (\theta_i - \overline{\theta}).$$

Thus, take a first-order Taylor series approximation around a point of $\theta_i = \theta$ for all i:

$$\theta^a \approx \left[\overline{\theta} + \sum_{i=1}^{N} \overline{\lambda}_i (\theta - \overline{\theta}) \right] + \sum_{i=1}^{N} \left[\frac{\partial \lambda_i}{\partial \theta_i} \theta_i + \lambda_i \right]_{\theta_i = \theta} (\theta_i - \theta)$$
$$= \theta + \sum_{i=1}^{N} \overline{\lambda}_i (\theta_i - \theta).$$

Where we use the fact that

$$\sum_{i=1}^{N} \left[\frac{\partial \lambda_i}{\partial \theta_i} \theta_i \right]_{\theta_i = \theta} = 0,$$

$$\sum_{i=1}^{N} \overline{\lambda}_i = 1,$$

and the definition of $\overline{\lambda}_i$:

$$\overline{\lambda}_i \equiv [\lambda_i]_{\theta_i = \theta}.$$

2.1 Case 1: $\alpha_i = \alpha$*,* $\sigma_i^2 = \sigma^2$*, and* $\sigma_{ij} = \rho\sigma^2$

If $\alpha_i = \alpha$, $\sigma_i^2 = \sigma^2$, and $\sigma_{ij} = \rho\sigma^2$, $\overline{\lambda}_i$ can be written as

$$\overline{\lambda}_i = \frac{\frac{\alpha^2\sigma^2}{1-\theta^2} + \sum_{j=1, j\neq i}^{N} \frac{\alpha^2\rho\sigma^2}{1-\theta^2}}{\sum_{i=1}^{N} \frac{\alpha^2\sigma^2}{1-\theta^2} + \sum_{i=1}^{N} \sum_{j=1, j\neq i}^{N} \frac{\alpha^2\rho\sigma^2}{1-\theta^2}}$$
$$= \frac{\frac{\alpha^2\sigma^2}{1-\theta^2} + \sum_{j=1, j\neq i}^{N} \frac{\alpha^2\rho\sigma^2}{1-\theta^2}}{N\left(\frac{\alpha^2\sigma^2}{1-\theta^2} + \sum_{j=1, j\neq i}^{N} \frac{\alpha^2\rho\sigma^2}{1-\theta^2} \right)}$$
$$= \frac{1}{N}.$$

Thus, as claimed,

$$\theta^a = \theta + \sum_{i=1}^{N} \frac{1}{N}(\theta_i - \theta) = \frac{1}{N}\sum_{i=1}^{N}\theta_i \ .$$

2.2 Case 2: $\sigma_i^2 = \sigma_{ij} = \sigma^2$

If $\sigma_i^2 = \sigma_{ij} = \sigma^2$, $\overline{\lambda}_i$ is

$$\overline{\lambda}_i = \frac{\dfrac{\alpha_i^2\sigma^2}{1-\theta^2} + \sum_{j=1,\,j\neq i}^{N} \dfrac{\alpha_i\alpha_j\sigma^2}{1-\theta^2}}{\sum_{i=1}^{N}\dfrac{\alpha_i^2\sigma^2}{1-\theta^2} + \sum_{i=1}^{N}\sum_{j=1,\,j\neq i}^{N}\dfrac{\alpha_i\alpha_j\sigma^2}{1-\theta^2}}$$

$$= \frac{\alpha_i^2 + \sum_{j=1,\,j\neq i}^{N}\alpha_i\alpha_j}{\sum_{i=1}^{N}\alpha_i^2 + \sum_{i=1}^{N}\sum_{j=1,\,j\neq i}^{N}\alpha_i\alpha_j}$$

$$= \frac{\alpha_i \sum_{i=1}^{N}\alpha_i}{\sum_{i=1}^{N}\alpha_i}$$

$$= \alpha_i .$$

Thus,

$$\theta^a = \theta + \sum_{i=1}^{N}\alpha_i(\theta_i - \theta) = \sum_{i=1}^{N}\alpha_i\theta_i = \overline{\theta} \ .$$

APPENDIX B

B.1 Group Mean Estimators

Consider the following heterogeneous dynamic model of disaggregated real exchange rate:

$$q_{it} = \alpha_i + \sum_{k=1}^{K}\rho_{ik}q_{it-k} + \nu_{it} \ , \ \nu_{it} \sim (0, \sigma_i^2)$$

$$\rho_{ik} = \rho_k + \xi_{1i} ,$$

$$\alpha_i = \alpha + \xi_{2i} .$$

We can rewrite the model as

$$q_j = X_j \beta_j + \nu_j$$

where $q_j = vec(q_{j,1}, q_{j,2}, ..., q_{j,T})$, $X_j = \begin{pmatrix} 1 & q_{j,0} & q_{j,-1} & \cdots & q_{j,-k} \\ 1 & q_{j,1} & q_{j,0} & \cdots & q_{j,-k+1} \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ 1 & q_{j,T-2} & q_{j,T-3} & \cdots & q_{j,T-k+1} \\ 1 & q_{j,T-1} & q_{j,T-2} & \cdots & q_{j,T-k} \end{pmatrix}$,

and $\nu_j = vec(\nu_{j,1}, \nu_{j,2}, ..., \nu_{j,T})$.

Assume that

$$\beta_j = \beta + \tau_j,$$

and

$$E(\tau_j) = 0,\ E(\tau_j \tau_j^{'}) = \Gamma.$$

Let B_j be the j-th ordinary least squares coefficient vector, the *mean group estimator* $\hat{\beta}_{MG}$ is simply given by

$$\hat{\beta}_{MG} = \sum_{j=1}^{N} \mathrm{B}_j .$$

Further, the *random coefficient estimator* $\hat{\beta}_{RCM}$ is given by

$$\hat{\beta}_{RCM} = \sum_{j=1}^{N} W_j \mathrm{B}_j ,$$

where

$$W_j = \left[\sum_{s=1}^{N} (\Gamma + V_s)^{-1} \right]^{-1} (\Gamma + V_j)^{-1},$$

$$V_j = \sigma_j^2 (X_j^{'} X_j)^{-1}.$$

B.2 Bias Correcting Methods

B.2.1 Kilian (1998)

Given a covariance stationary AR(p) process,

$$y_t = \alpha + \beta_1 y_{t-1} + \beta_2 y_{t-2} + \cdots + \beta_p y_{t-p} + u_t,$$

where $u_t \sim^{i.i.d.} (0, \Sigma_u)$.

Step 1a: Estimate the AR(p) model and obtain the estimates $\hat{\mathrm{B}} = vec(\hat{\beta}_1, \hat{\beta}_2, \cdots \hat{\beta}_p)$. Generate 1000 bootstrap replications $\hat{\mathrm{B}}^*$ from

$$y_t^* = \hat{\alpha} + \hat{\beta}_1 y_{t-1}^* + \hat{\beta}_2 y_{t-2}^* + \cdots + \hat{\beta}_p y_{t-p}^* + u_t^*, \qquad (*)$$

using standard nonparametric bootstrap procedures. Then approximate the bias term $\Psi = E(\hat{\mathrm{B}} - \mathrm{B})$ by $\Psi^* = E^*(\hat{\mathrm{B}}^* - \hat{\mathrm{B}})$; this suggests the bias estimate $\hat{\Psi} = \overline{\hat{\mathrm{B}}}^* - \hat{\mathrm{B}}$

Step 1b: Calculate the modulus of the largest root of the companion matrix associated with $\hat{\mathrm{B}}$ as $m(\hat{\mathrm{B}})$. If $m(\hat{\mathrm{B}}) \geq 1$, set $\tilde{\mathrm{B}} = \hat{\mathrm{B}}$. If $m(\hat{\mathrm{B}}) < 1$, construct the bias-corrected coefficient estimate $\tilde{\mathrm{B}} = \hat{\mathrm{B}} - \hat{\Psi}$. Plug $\tilde{\mathrm{B}}$ into $m(\cdot)$. If $m(\tilde{\mathrm{B}}) \geq 1$, let $\hat{\Psi}_1 = \hat{\Psi}$ and $\delta_1 = 1$. Define $\hat{\Psi}_{i+1} = \delta_i \hat{\Psi}_i$ and $\delta_{i+1} = \delta_i - 0.01$. Set $\tilde{\mathrm{B}} = \tilde{\mathrm{B}}_i$ after iterating on $\tilde{\mathrm{B}}_i = \hat{\mathrm{B}} - \hat{\Psi}_i, i = 1, 2, \ldots,$ until $m(\tilde{\mathrm{B}}_i) < 1$.

Step 2a: Substitute $\tilde{\mathrm{B}}$ for $\hat{\mathrm{B}}$ in equation (*) and generate 2000 new bootstrap replications $\hat{\mathrm{B}}^*$ using standard nonparametric bootstrap procedures. In order to reduce the computational requirements, we follow Kilian's (1998) suggestion that use the first-stage bias estimate $\hat{\Psi}$ as a proxy for $\hat{\Psi}^*$.

Step 2b: Calculate $\tilde{B}^*$ from $\hat{B}^*$ and $\hat{\Psi}^*$, following the same procedures in step 1b with obvious changes in notation.

Step 3: Calculate the 25 and 75 percentile interval endpoints of the distribution of impulse response function.

B.2.2 So and Shin (1999)

Consider the following simple AR(1) model as an example, the extension to higher order AR(p) model is straightforward. Assume

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

the bias-corrected estimate of ρ by recursive demeaning method proposed by So and Shin (1999) is given by

$$\hat{\rho}_{RE} = \frac{\sum_{t=2}^{T}(y_{t-1} - \bar{y}_{t-1})(y_t - \bar{y}_{t-1})}{\sum_{t=2}^{T}(y_{t-1} - \bar{y}_{t-1})^2},$$

where

$$\bar{y}_{t-1} = \frac{1}{t-1}\sum_{s=1}^{t-1} y_s .$$

APPENDIX C

Table C1: Data Source. Price data for European countries is obtained from Eurostat. Most U.S. price data comes from BLS with some exceptions collected from Eurostat due to unavailability. Codes of data are provided in brackets.

	Variable	European Countries	U.S.
1	BC	Bread and Cereals (1111)	CEREALS & BAKERY PRODUCTS (USCPFCB.F)
2	Meat	Meat (1112)	MEATS (USCPFAM.F)
3	Dairy	Dairy Products (1114)	DAIRY & RELATED PRODUCTS (USCPFDY.F)
4	Fru	Fruit (1116)	FRESH FRUITS (USCPFFF.F)
5	Alco	Alcoholic and non-alcoholic drinks (1150)	0.57*(ALCOHOLIC BEVERAGES)+0.43*(NONALCOHOLIC BEVERAGES & BEVERAGE MATERIALS) a) ALCOHOLIC BEVERAGES (USCPBAL.E) b) NONALCOHOLIC BEVERAGES & BEVERAGE MATERIALS (USCPBNA.E)
6	Clot	Clothing (1210)	APPAREL (USCPAPPLE)
7	Foot	Footwear (1220)	FOOTWEAR (USCPAS..E)
8	Fuel	Fuel and Energy (1330)	FUEL & UTILITIES (USCPHUT.E)
9	Furn	Furniture (1410)	FURNITURE & BEDDING (USCPHGF.F)
10	Dapp	Domestic Appliances (1420)	Domestic Appliances (1420) – Eurostat
11	Vehl	Vehicles (1610)	NEW VEHICLES (USCPTVN.E)
12	Ptran	Public Transport (1630)	PUBLIC TRANSPORTATION (USCPTB..F)
13	Sound	Sound and photo equipment (1710)	Sound and photo equipment (1710) – Eurostat
14	Book	Books (1730)	Books (1730) – Eurostat
15	Lei	Lei (1720)	Lei (1720) – Eurostat
16	Hotl	Hotels (1830)	LODGING AWAY FROM HOME INCLUDING HOTELS (USCPHST.F)

Table C2: Data Availability for Each Country

	Belgium	Denmark	Germany	Spain	France
Bread and Cereals (1111)	81:01-96:11	81:01-96:09	81:01-95:07	81:01-96:10	81:01-96:10
Meat (1112)	81:01-96:11	81:01-96:09	81:01-95:07	81:01-96:10	81:01-96:10
Dairy Products (1114)	81:01-96:11	81:01-96:09	81:01-95:07	81:01-96:10	81:01-96:10
Fruit (1116)	81:01-96:11	81:01-96:09	81:01-95:07	81:01-96:10	81:01-96:10
Alcoholic and non-alcoholic drinks (1150)	81:01-96:11	NA	81:01-95:07	81:01-96:10	81:01-96:10
Clothing (1210)	81:01-96:11	NA	81:01-95:07	81:01-96:10	81:01-96:10
Footwear (1220)	81:01-96:11	81:01-96:09	81:01-95:07	81:01-96:10	81:01-96:10
Fuel and Energy (1330)	81:01-96:03	81:01-95:10	81:01-95:07	81:01-96:10	81:01-96:10
Furniture (1410)	81:01-94:10	81:01-94:10	81:01-94:10	81:01-94:10	81:01-94:10
Domestic Appliances (1420)	81:01-94:10	NA	81:01-94:10	81:01-94:10	81:01-94:10
Vehicles (1610)	81:01-96:11	NA	81:01-95:07	81:01-96:10	81:01-96:10
Public Transport (1630)	NA	82:04-96:09	82:04-95:07	82:04-96:10	82:04-96:10
Sound and photo equipment (1710)	81:01-96:11	NA	81:01-95:07	81:01-96:10	81:01-96:10
Leisure (1720)	81:01-96:11	NA	NA	81:01-96:10	81:01-96:10
Books (1730)	81:01-96:11	81:01-96:09	81:01-95:07	NA	81:01-96:10
Hotels (1830)	81:01-96:11	81:01-96:09	81:01-95:07	81:01-96:10	81:01-96:10
Total # of sectors	15	10	15	15	16

	Italy	Netherlands	Portugal	U.K.	U.S.
Bread and Cereals (1111)	81:01-95:12	NA	81:01-96:02	81:01-96:10	81:01-96:12
Meat (1112)	81:01-95:12	81:01-96:10	81:01-96:02	81:01-96:10	81:01-96:12
Dairy Products (1114)	81:01-95:12	81:01-96:10	NA	81:01-96:10	81:01-96:12
Fruit (1116)	81:01-95:12	81:01-96:10	NA	81:01-96:10	81:01-96:12
Alcoholic and non-alcoholic drinks (1150)	81:01-95:12	81:01-96:10	81:01-95:11	81:01-96:10	81:01-96:12
Clothing (1210)	81:01-95:12	81:01-96:10	81:01-96:02	81:01-96:10	81:01-96:12
Footwear (1220)	81:01-95:12	81:01-96:10	81:01-96:02	81:01-96:10	81:01-96:12
Fuel and Energy (1330)	81:01-95:12	81:01-96:10	81:01-96:02	81:01-96:10	81:01-96:12
Furniture (1410)	81:01-94:09	81:01-94:10	81:01-94:10	81:01-94:10	81:01-96:12
Domestic Appliances (1420)	81:01-94:09	81:01-94:10	81:01-94:10	81:01-94:10	81:01-94:10
Vehicles (1610)	81:01-95:12	81:01-96:10	NA	81:01-96:10	81:01-96:12
Public Transport (1630)	82:04-95:12	82:04-96:10	NA	82:04-96:10	81:01-96:12
Sound and photo equipment (1710)	81:01-95:12	81:01-96:10	NA	81:01-96:10	81:01-95:09
Leisure (1720)	NA	81:01-96:10	81:01-96:02	81:01-96:10	81:01-95:09
Books (1730)	81:01-95:12	81:01-96:10	NA	81:01-96:10	81:01-95:09
Hotels (1830)	81:01-95:12	81:01-96:10	81:01-96:02	81:01-96:10	81:01-96:12
Total # of sectors	15	15	10	16	16

APPENDIX D

Figure 1: Prices: Imbs et al. (2002) vs. Corrected Data

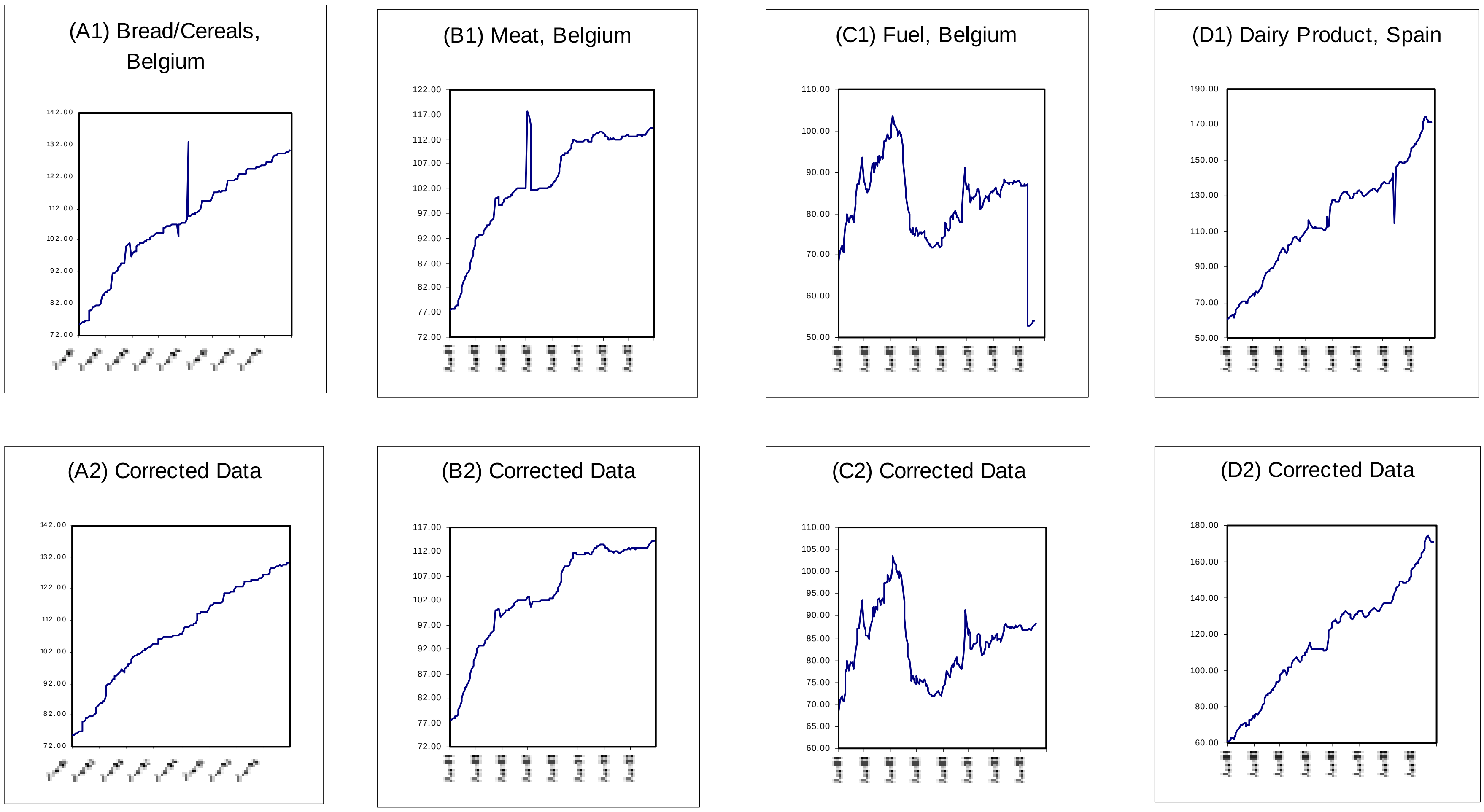

Figure 2: Prices: Imbs et al. (2002) vs. Corrected Data (continued)