PPP STRIKES BACK: AGGREGATION AND THE REAL EXCHANGE RATE*

Jean Imbs

Haroon Mumtaz

Morten O. Ravn

Helene Rey

Abstract

We show the importance of a dynamic aggregation bias in accounting for the PPP puzzle. We prove that the aggregate real exchange rate is persistent because its components have heterogeneous dynamics. Established time series and panel methods fail to control for this. Using Eurostat data, we find that when heterogeneity is taken into account, the estimated persistence of real exchange rates falls dramatically. Its half-life, for instance, may fall to as low as eleven months, significantly below the 'consensus view' of three to five years, summarized in Rogoff [1996].

* *Forthcoming Quarterly Journal of Economics.* We thank the editor, two referees, Pól Antras, Michael Devereux, Charles Engel, Martin Evans, Lutz Kilian, M. Hashem Pesaran, Richard Portes, Roberto Rigobon, Kenneth Rogoff, Barbara Rossi, Christopher Sims, Ron Smith, Antonio Spilimbergo, Shang-Jin Wei, Michael Woodford, Charles Wyplosz, and participants at numerous conferences and seminars for helpful comments. Parts of this paper were completed while Ravn was a visitor at the University of Pennsylvania, and Imbs and Rey were visitors at the International Monetary Fund. The authors are grateful to these institutions for their hospitality. The paper is part of the project "Exchange Rates, International Relative Prices, and Macroeconomic Models", funded by the ESRC (grant no.L138 25 1043), and of the RTN programme "The Analysis of International Capital Markets: Understanding Europe's Role in the Global Economy", funded by the European Commission (contract no. HPRN-CT-1999-00067). This paper does not represent the views of the Bank of England or of Monetary Policy Committee members. The work was done while the second author was at London Business School.

I. Introduction

The study of real exchange rates, defined as the international relative price of a basket of goods expressed in a common currency, is perhaps the most intensely researched area in international macroeconomics. Absolute purchasing power parity states that real exchange rates should be constant and equal to one or, expressed in relative terms, that there should be no changes in the real exchange rate. Yet the consensus emerging from an extensive literature appears to be that, although real exchange rates may converge to parity in the long run, the rate at which this happens is very slow. The speed of mean reversion is usually summarized by the half-life, the time necessary for half the effect of a given shock to dissipate. According to Rogoff's [1996] authoritative survey, standard estimates of the real exchange rate half-life lay in the range of three to five years. While the high volatility of real exchange rates could potentially be explained by monetary (or financial) shocks, the rate of mean reversion appears too slow to be compatible with plausible nominal rigidities. Hence, Rogoff argues, the "PPP puzzle".[1] Evidence on the law of one price (LOP) is hardly more encouraging, as it suggests persistent international differences in goods prices as well.[2]

This paper takes issue with the "consensus view" and shows that slow mean reversion in the aggregate real exchange rate is consistent with - on average - much faster adjustment of disaggregated relative prices. Existing estimates of real exchange rate persistence are based upon the (implicit) assumption that all relative prices composing the real exchange rate converge to parity at the same speed. But there is little (if any) theoretical justification for this assumption; indeed, it is hard to think of reasons why clothes and vegetables, say, should revert to parity at the same speed.

We show how the failure to allow for heterogeneity in price adjustment dynamics at the good level induces a *positive* bias in persistence estimates, irrespective whether the estimation is performed using pure time series, a panel of aggregate real exchange rates, or even a panel of sectoral relative prices. We stress the importance of correcting for heterogeneity when estimating persistence in the real exchange rate. When heterogeneity in adjustment dynamics is allowed for, panel data estimates point to an average speed of mean reversion faster than the consensus view. The persistence of disaggregated relative prices is on average smaller than the persistence of the aggregate real exchange rate itself.[3] *Importantly, our result does not require nor imply that*

[1] See Froot and Rogoff [1995] or Coakley, Flood and Taylor [2002] for other surveys. Recently, studies emphasizing non-linearities argue that the true half life is smaller than the consensus estimate, for instance in Taylor, Peel and Sarno [2001]. Others argue instead that it could in fact be much bigger (Murray and Papell [2002c]) or that the confidence intervals are far too wide to tell (Murray and Papell [2002a], Rossi [2003], Kilian and Zha [2002]).

[2] See Goldberg and Knetter [1997] for a survey. Classic studies include Giovannini [1988], Isard [1977], Knetter [1989, 1993], and Richardson [1978]. Recent studies include Crucini, Telmer and Zachariadis [2001], Haskel and Wolf [2001] or Parsley and Wei [2003].

[3] Going back to Isard [1977], an extensive literature has examined the behavior of disaggregated relative prices, with focus on estimates of the exchange rate pass-through. For a recent study see Engel [2000b].

persistence be systematically lower at the disaggregated level.[4] If persistence were homogeneously low across all disaggregated relative prices, there would be no PPP puzzle in the first place because aggregate persistence would also appear low.[5]

Our contention is not that standard estimates of real exchange rate persistence are "wrong", as the notion of a bias might suggest. The real exchange rate is a well-defined object, and one can study its properties using standard techniques. Rather, we take issue with the interpretation of the standard results. The PPP puzzle arises because the estimated real exchange rate persistence is construed to be excessive in reference to theories where differences in prices are sustained by limits to arbitrage or nominal rigidities. Here we argue that impediments to arbitrage or nominal rigidities have every reason to vary with each good's characteristics.[6] It is this heterogeneity that we find to be an important determinant of the observed real exchange rate persistence, since it gives rise to highly persistent *aggregate* series while relative price persistence is low on average at a *disaggregated* level.

We quantify the bias using an international sectoral price database issued by Eurostat. We find it to be substantial in these data. Our preferred estimate of the half life is eleven months with a confidence interval ranging from seven to twelve months. This is far below standard estimates, and it is not due to any specificities in our data, since we reproduce the 'consensus view' when we do not correct for heterogeneity.

Our results appear to be robust. First, we consider the potential impact of measurement error that may give rise to an attenuating bias in the sectoral autoregressive parameters. Were this the case, however, persistence estimates would be systematically lower at the sectoral level - which they are not. Formal tests for errors-in-variables also provide no evidence supporting the presence of measurement error. Furthermore, we confirm our results in other versions of the Eurostat dataset.[7] Second, a recent strand of literature argues that, when the underlying data generating process is highly persistent, small sample least squares estimates of persistence tend to be biased downwards. We implement a bias reduction method on our disaggregated data. Our bias corrected half-life estimate rises from eleven to eighteen months only and is estimated precisely with a confidence interval that continues to exclude the 'consensus view' range of three to five years.[8]

The remainder of the paper is structured as follows. We next describe in detail the bias that plagues dynamic panel and time series estimates when there is sectoral heterogeneity. We

---

[4] In that sense, our results are different from - but not in contradiction with - Parsley and Wei [1996], and Crucini and Shintani [2002], who find *homogeneously* rapid reversion to parity.

[5] Other biases may be relevant as well. Taylor [2001], for example, studies *temporal* aggregation issues. Temporal and sectoral aggregation biases are distinct conceptually and may well both be present at the same time.

[6] For instance, Blanchard [1987] or more recently Bils and Klenow [2002] discuss the relevance of heterogeneity in price adjustments at the disaggregated level.

[7] For more details, see Imbs, Mumtaz, Ravn and Rey [2004].

[8] In Imbs, Mumtaz, Ravn and Rey [2004], we detail the reasons why the small sample bias is limited in our dataset.

derive conditions under which the bias is positive and show it increases with the extent of heterogeneity. Section III reviews various existing procedures used to estimate half-lives and presents the estimator we implement to control for heterogeneity. Section IV introduces the data, performs basic tests and shows that the conditions for the positivity of the bias are borne out in the data. In Section V we first reproduce standard results, then test for heterogeneity and find strong support for heterogeneous dynamics across sectors. Accordingly, we use estimators that allow for heterogeneity. Persistence drops dramatically. Section VI examines alternative explanations for our findings and performs robustness checks. Section VII concludes.

II. Heterogeneous Adjustment Dynamics in Theory

This section explains how the failure to account for heterogeneity in relative price dynamics gives rise to a positive bias whose magnitude increases in the degree of heterogeneity. We do this in three steps. First, we focus on a panel of sectoral relative prices, and show the conditions for an upward bias that rises with heterogeneity.[9] Second, we show that the problem subsists in (time-series) estimations using aggregate real exchange rates. Third, we extend the result to panels of real exchange rates.

II.A. Bias in Dynamic Heterogeneous Panels

We build on the work of Pesaran and Smith [1995], who generalize the insights of Robertson and Symons [1992] on the econometric issues arising in panels with heterogeneous dynamics. We show the conditions under which standard estimators will be biased upwards in panels of sectoral real exchange rates.

Consider estimating the mean persistence of sectoral real exchange rates in a panel consisting of N cross-sectional units. To simplify, but without loss of generality, assume that each of the panel elements is given by a first-order autoregressive process:

$$q_{it} = \gamma_i + \rho_i q_{it-1} + \varepsilon_{it}, \;\; i = 1, ..., N, \quad t = 1, ..., T \tag{1}$$

with $\gamma_i = \gamma + \eta_i^{\gamma}$ and $\rho_i = \rho + \eta_i^{\rho}$. We assume that η_i^{γ} and η_i^{ρ} have zero mean and constant covariance, and that the set of random coefficients ρ_i has support on the interval $]-1, 1[$. Furthermore, ε_{it} is assumed to be independently distributed with mean 0 and variance σ_i^2.[10] We seek to estimate ρ, the average persistence of the relative prices. We order the N sectors so that for all i, $\rho_{i+1} \geq \rho_i$, $\rho_i \in (0, 1)$, where we impose, realistically, that relative prices are positively serially correlated.

[9] Throughout the paper we use the terms sectoral relative prices and sectoral real exchange rates interchangeably.

[10] This section follows closely Pesaran and Smith [1995] including their distributional assumptions, and is used to build the intuition for our results. We relax the assumption of zero cross sectoral correlation in the next section.

An estimation where the persistence parameters are constrained to be homogeneous across sectors would have the following form:

$$\begin{aligned} q_{it} &= \gamma_i + \rho q_{it-1} + e_{it} && (2) \\ e_{it} &= \eta_i^{\rho} q_{it-1} + \varepsilon_{it} && (3) \end{aligned}$$

It follows immediately that as soon as the dynamics of the panel units are constrained to be homogeneous, the lagged dependent variable enters the error term and estimates of ρ are inconsistent.

Now let ρ^Q denote the fixed effect estimator of the first order autoregressive coefficient. Pesaran and Smith [1995] show that

$$\text{plim}_{N\to\infty,T\to\infty}\ \left(\rho^Q\right) = \rho + \Delta$$

where

$$\Delta = \frac{1}{N}\sum_{i=1}^{N}\left(\frac{\eta_i^{\rho}\sigma_i^2}{1-\rho_i^2}\right) / \left[\frac{1}{N}\sum_{i=1}^{N}\left(\frac{\sigma_i^2}{1-\rho_i^2}\right)\right] = \sum_{i=1}^{N}\left(\rho_i-\rho\right)\frac{\sigma_i^2}{1-\rho_i^2} / \left[\sum_{i=1}^{N}\left(\frac{\sigma_i^2}{1-\rho_i^2}\right)\right].$$

Hence the expression of the bias is given by

$$\begin{aligned} \Delta &= \sum_{i=1}^{N}\left(\rho_i-\rho\right)\alpha_i, \\ \alpha_i &= \frac{\sigma_i^2}{1-\rho_i^2} / \left[\sum_{i=1}^{N}\left(\frac{\sigma_i^2}{1-\rho_i^2}\right)\right]. \end{aligned}$$

Proposition 1. The sign of the bias arising from the failure to account for dynamic heterogeneity across panel units is given by the sign of $\Delta = \sum_{i=1}^{N}\left(\rho_i-\rho\right)\alpha_i$. For a large N, the bias is therefore positive *if and only if* $cov\left(\widetilde{\rho},\widetilde{\alpha}\right) > 0$, i.e. the covariance between the vector of persistence parameters $\widetilde{\rho} = \{\rho_i\}_{i=1}^{N}$ and the vector of coefficients $\widetilde{\alpha} = \{\alpha_i\}_{i=1}^{N}$ is positive.

Proof of Proposition 1. From the above derivation, it is immediate that the sign of the bias is the same as the sign of Δ. By definition, $cov\left(\widetilde{\rho},\widetilde{\alpha}\right) = \lim\limits_{N\to+\infty} 1/(N(N-1))\sum_{i=1}^{N}\left(\rho_i-\rho\right)\left(\alpha_i-\alpha\right)$ where α is the mean of the coefficients $\{\alpha_i\}_{i=1}^{N}$.

By definition of the ρ_i's $\lim\limits_{N\to+\infty} 1/(N(N-1))\sum_{i=1}^{N}\left(\rho_i-\rho\right)\alpha = 0$. The sign of $cov\left(\widetilde{\rho},\widetilde{\alpha}\right)$ is therefore the same as the sign of $1/(N(N-1))\sum_{i=1}^{N}\left(\rho_i-\rho\right)\alpha_i$, which is the sign of Δ.

Corollary 1.1. A sufficient condition for the dynamic heterogeneity bias to be positive is that $0 \leq \alpha_i \leq \alpha_{i+1}$ for all i.

Proof of Corollary 1.1. See Appendix 1.

Section IV verifies that the condition specified in Proposition 1 is borne out in our data. The intuition for the sign of the bias can be understood straightforwardly with the help of the sufficient condition described in Corollary 1.1. If $0 \leq \alpha_i \leq \alpha_{i+1}$, then α_i is higher for large realizations of $\rho_i - \rho$, and the fixed effects estimates of ρ^Q are dominated by the components of the relative prices that revert to parity the slowest. In practice it is enough, when N is large, that α_i *tend to be higher* for large realizations of $\rho_i - \rho$. This is equivalent to checking that $cov(\widetilde{\rho}, \widetilde{\alpha}) > 0$ holds in the data.

Corollary 1.2. A (positive) bias tends to increase, ceteris paribus, with the cross-sectoral dispersion in persistence.

Proof of Corollary 1.2. See Appendix 1.

It is also evident that any type of instrumentation will not solve the problem: any instrument highly correlated with the dependent variable will unavoidably also be correlated with the error term. In the presence of a lagged dependent variable, a common approach to handling the presence of fixed effects is to first-difference the data and use the IV or GMM estimators suggested in Anderson and Hsiao [1982] and Arellano and Bond [1991]. But under dynamic heterogeneity, this will still lead to inconsistent estimates since

$$\begin{aligned} \triangle q_{it} &= \rho \triangle q_{it-1} + \triangle e_{it} \\ \triangle e_{it} &= \eta_i^\rho \triangle q_{it-1} + \triangle \varepsilon_{it}. \end{aligned}$$

Standard panel data estimators suffer from inconsistency when there is dynamic heterogeneity across panel units and, under plausible conditions, they will overestimate the average persistence of relative prices.

The vast majority of papers dealing with the PPP puzzle base their estimates of relative price persistence not on sectoral data but on time series (or panels) of aggregate real exchange rates. We next show why aggregation fails to solve the problem created by heterogeneity.

II.B. Aggregation Bias: Time Series

This section describes how heterogeneous dynamics at the sectoral level translates into biased aggregate estimates. We first focus on the case where the panel consists of the relative prices of goods for a single country pair. Consider an economy with N sectors indexed by i and assume as above that:

$$q_{it} = \gamma_i + \rho_i \, q_{it-1} + \varepsilon_{it}, \;\; i = 1, ..., N, \quad t = 1, ..., T$$

with $\gamma_i = \gamma + \eta_i^\gamma$ and $\rho_i = \rho + \eta_i^\rho$. We now allow for non-zero cross-sectoral covariances of ε_{it}, with $E(\varepsilon_{it} \, \varepsilon_{jt}) = \sigma_{ij}$ for $i \neq j$. These correlations could arise, for example, from common shocks across goods or from omitted (unobservable) global influences. We again order the sectors so that ρ_i is non-decreasing in i, positive and strictly less than unity. It follows that

$$\begin{aligned}\sigma_{q_i}^2 &= \frac{\sigma_i^2}{1-\rho_i^2} \\ \sigma_{q_i,q_j} &= \frac{\sigma_{ij}}{1-\rho_i\rho_j}.\end{aligned}$$

where $\sigma_{q_i}^2$ is the variance of q_i and σ_{q_i,q_j} is the covariance between q_i and q_j. The bilateral real exchange rate q_t can be approximated by a linear aggregation of the different sectors with weights ω_j associated with the jth good[11]

$$q_t = \sum_{j=1}^{N} \omega_j q_{jt}, \quad \sum_{j=1}^{N} \omega_j = 1.$$

In general, q_t can be written as

$$\left[\prod_{i=1}^{N} (1-\rho_i L)\right] q_t = \sum_{i=1}^{N} \omega_i \left(\prod_{j\neq i} (1-\rho_j L)\right) \varepsilon_{it}.$$

where L denotes the lag operator. As is well-known, cross-sectional aggregation of N $AR(1)$ processes produces an $ARMA(N, N-1)$.[12] If the dynamics of the cross-sectional units were homogeneous, $\rho_i = \rho$ for all $i = 1, .., N$, the roots would cancel out, and this ARMA process would simplify into an autoregressive process of order one.[13] Allowing for heterogeneity by simply estimating a large order ARMA process for the real exchange rate is a theoretical possibility. But pursuing this route will be impossible in most cases, for lack of degrees of freedom, unless the sample period is long enough.[14] Heterogeneous estimators are better-suited to tackling the issue than estimating processes with high order-ARMA terms.

In fact, the vast majority of PPP studies estimate the persistence of the real exchange assuming that its dynamics are best described by an AR(p) process. Many studies actually use an

[11] This is a log-linear approximation to the CPI- based real exchange rate when CPI weights are equal across countries (see Appendix A1 in Imbs, Mumtaz, Ravn and Rey [2004]).

[12] Granger and Morris [1976] show that an aggregate of N $ARMA(p_i, q_i)$ is an $ARMA(P, Q)$ process where $P \leq \sum_i p_i$ and $Q \leq \max_j (P - p_j + q_j)$. The size of P and Q will depend on the degree of heterogeneity of the dynamics of the underlying data. Granger (1980) shows further that if N $AR(1)$ series are aggregated and the autoregressive parameters can take on *any* value (as opposed to, say, M discrete values) in a given interval, the aggregated data will correspond to no $ARMA$ process with a finite number of parameters. He derives examples in which the aggregated data displays long memory.

[13] In this case $(1-\rho L)^N q_t = (1-\rho L)^{N-1} \sum_i \omega_i \varepsilon_{it}$ so that $q_t = \rho q_{t-1} + \sum_i \omega_i \varepsilon_{it}$.

[14] If sectoral relative prices follow AR(I) processes, the real exchange rate is an $ARMA[N{\cdot}I, (N-1){\cdot}I]$. Suppose $I = 2$ and $N = 200$, this implies estimating an $ARMA(400, 398)$. Pesaran and Smith [1995] recommend *infinite* distributed lag specifications. Few datasets (and certainly not ours) can afford this kind of degrees of freedom.

AR(1) as their standard specification.[15] So will we to simplify the derivations. We have

$$\begin{aligned} q_t &= \gamma + \rho q_{t-1} + \varepsilon_t \\ \gamma &= \sum_{j=1}^{N} \omega_j \gamma_j, \ \varepsilon_t = \sum_{j=1}^{N} \omega_j \varepsilon_{jt} + \sum_{j=1}^{N} \eta_j^{\rho} \omega_j q_{jt-1}. \end{aligned}$$

Thus, the lagged dependent variables are present in the error term and we can show as in the previous section that this 'aggregation bias' is positive under plausible conditions. To economize on notation but without loss of generality, we now assume that $\gamma = 0$. Consider the least squares estimate of the first-order autoregressive coefficient of q_t, given by $\rho^Q = E(q_t q_{t-1}) / E(q_t^2)$. We can derive (see Appendix 2) that

$$\text{plim}_{N\to\infty, T\to\infty} \left(\rho^Q\right) = \rho + \Delta$$

with

$$\Delta = (\rho_i - \rho) \left[\sum_{i=1}^{N} \frac{\omega_i^2}{1-\rho_i^2} \sigma_i^2 + \sum_{i \neq j}^{N} \frac{\omega_i \omega_j \sigma_{ij}}{1 - \rho_i \rho_j} \right] / \left[\sum_{i=1}^{N} \left(\frac{\omega_i^2}{1-\rho_i^2} \sigma_i^2 + \sum_{i \neq j}^{N} \frac{\omega_i \omega_j \sigma_{ij}}{1-\rho_i \rho_j} \right) \right].$$

It is useful to rewrite the bias as

$$\Delta = \sum_{i=1}^{N} (\rho_i - \rho) \delta_i,$$

with

$$\delta_i = \left[\frac{\omega_i^2 \sigma_i^2}{1-\rho_i^2} + \sum_{i \neq j}^{N} \frac{\omega_i \omega_j \sigma_{ij}}{1-\rho_i \rho_j} \right] / \left[\sum_{i=1}^{N} \left(\frac{\omega_i^2}{1-\rho_i^2} \sigma_i^2 + \sum_{i \neq j}^{N} \frac{\omega_i \omega_j \sigma_{ij}}{1-\rho_i \rho_j} \right) \right].$$

We can now spell out the conditions for the bias to be positive.

PROPOSITION 2. The sign of the bias arising from the failure to account for dynamic heterogeneity across real exchange rate components is given by the sign of $\Delta = \sum_{i=1}^{N} (\rho_i - \rho) \delta_i$. For a large N, the bias is therefore positive *if and only if* $cov\left(\widetilde{\rho}, \widetilde{\delta}\right) > 0$, i.e. the covariance between the vector of persistence parameters $\widetilde{\rho} = \{\rho_i\}_{i=1}^{N}$ and the vector of coefficients $\widetilde{\delta} = \{\delta_i\}_{i=1}^{N}$ is positive.

Proof of Proposition 2. The proof is the same as the proof of Proposition 1, replacing α_i by δ_i.

COROLLARY 2.1. A sufficient condition for the dynamic heterogeneity bias to be positive is that $0 \leq \delta_i \leq \delta_{i+1}$ for all i.

[15] For instance, Choi, Mark and Sul [2003] assume all along an AR(1), a choice they label "conventional", and that is followed among many others by Taylor [2001] or Murray and Papell [2002b].

Proof of Corollary 2.1. Observe that $\sum_{i=1}^{N} \delta_i = 1$. For the same reasons as in Corollary 1.1, $0 \leq \delta_i \leq \delta_{i+1}$ for all i constitute a strong sufficient condition that ensures the positivity of Δ.

Corollary 2.2. A (positive) bias tends to increase *ceteris paribus* with the cross-sectoral dispersion in persistence.

Proof of Corollary 2.2. The proof is identical to that of corollary 1.2. since again the magnitude of Δ increases with the distance between sectoral persistence and its cross-sectional average.

We note the similarity between these results and the results for disaggregated panel estimators that we derived in the previous section. The same intuition carries through. Again we verify in section IV that the condition spelled out in Proposition 2 holds in our data, i.e. that the covariance between the estimation weights $\{\delta_i\}_{i=1}^{N}$ and the persistence parameters $\{\rho_i\}_{i=1}^{N}$ is positive.

Three points are worth stressing at this stage. First, Proposition 2 suggests the covariances between sectoral price residuals affect both the magnitude and sign of the bias. Hence, controlling for cross-sectoral correlations will be important in our empirical application. Second, Proposition 2 clarifies the role of the weights used in aggregating sectoral prices into the Consumer Price Index. While δ_i depends on ω_i, the sign or the magnitude of the relation is by no means straightforward. As will become clear, our heterogeneous estimators aggregate autoregressive coefficient estimates, not price series. The weight each good receives in the Consumer Price Index is only one of the determinants of the relation between δ_i and ρ_i. Third, our result does *not* imply that the persistence of the real exchange rate is not informative in itself. We merely stress that, in the face of heterogeneity, the persistence of the real exchange rate will not be a consistent estimate of the mean persistence of relative prices, ρ.

A simple example of conditions under which Corollary 2.1 is satisfied is when the CPI weights are similar, the innovation variances are similar and the covariances between the innovations are similar and positive (see Imbs, Mumtaz, Ravn and Rey [2004] for details).[16] We note that such families of restrictions (broadly defined) are plausible for sectoral price data. In section IV we show unambiguously that the coefficients δ_i and the persistence parameters $(\rho_i - \rho)$ *covary positively in our data*.

II.C. Aggregation Bias: Panels of Real Exchange Rates

Most recent papers study panels of aggregate real exchange rates. We now show that the insights developed above apply also in such panels. As is standard, we control for country fixed effects. We let the autoregressive coefficient vary *across sectors*, while the intercept is allowed to vary *across countries*.[17] The proof can easily be generalized to allow for heterogeneity in the

[16] Positive covariances rules out degenerate cases where some sectoral processes exactly cancel out.

[17] Country fixed effects are well-known to be important in real exchange rate estimation. See Frankel and Rose [1996]. It is straightforward, though not insightful, to generalize to an intercept varying across both countries and sectors.

autoregressive coefficients across countries, but one dimension of heterogeneity is sufficient for providing the basic insights. Our sectoral relative prices can be written as

$$q_{ict} = \gamma_c + \rho_i \; q_{ict-1} + \varepsilon_{ict}$$

where c denotes a given country. The fixed effect estimate of the first order autoregressive coefficient for the aggregate real exchange rate is given by

$$\text{plim}_{N\to\infty,T\to\infty} \left(\rho^Q_{FE}\right) = E\left[\left(Q_{ct} - \bar{Q}_c\right)\left(Q_{ct-1} - \bar{Q}_c\right)\right] / E\left(Q_{ct} - \bar{Q}_c\right)^2$$

where $\bar{Q}_c = E(Q_{ct}) = 1/N \sum_i \gamma_{ic}/\left(1-\rho_i\right)$ denotes a country-specific average level of the real exchange rate. Let $\tilde{q}_{ict} = q_{ict} - \gamma_{ic}/\left(1-\rho_i\right)$. It is immediate that $Q_{ct} - \bar{Q}_c = 1/N \sum_i \tilde{q}_{ict}$. Thus, since $\tilde{q}_{ict} = \rho_i \; \tilde{q}_{ict-1} + \varepsilon_{ict}$ by definition, we have

$$\text{plim}_{N\to\infty,T\to\infty} \left(\rho^Q_{FE}\right) = \frac{\sum_{i=1}^N \rho_i \sigma^2_{\tilde{q}_{ic}} + \sum_{i<j}^N \left(\rho_i + \rho_j\right) \sigma_{\tilde{q}_{ic},\tilde{q}_{jc}}}{\sum_{i=1}^N \sigma^2_{\tilde{q}_{ic}} + 2\sum_{i<j}^N \sigma_{\tilde{q}_{ic},\tilde{q}_{jc}}}$$

with obvious notations. Thus, the rest of the proof in the previous section carries through almost identically. A fixed effect estimator with country-specific intercepts continues to suffer from a positive bias under similar conditions.[18] We note that a test may reject heterogeneity in panels of *aggregate* real exchange rates if differences in relative price dynamics are mostly relevant at the *goods level*.

III. Econometric Methods

We measure persistence using three alternative statistics. In the PPP literature, the most commonly used measure is by far the 'half-life', denoted $T_{1/2}$ here, defined as the number of periods it takes until half the effect of a given shock dissipates. In the case of an autoregressive process of order one, $T_{1/2}$ can be computed as $\ln(0.5)/\ln(\widehat{\rho})$ where $\widehat{\rho}$ is the estimated first-order autoregressive coefficient. For higher order autoregressive models, we use the estimated impulse response function. We follow Kilian and Zha [2002] - among others - and define the half-life as the largest value of $T_{1/2}$ such that $\widehat{IR}\left(T_{1/2}-1\right) \geq 0.5$ and $\widehat{IR}\left(T_{1/2}\right) < 0.5$ where $\widehat{IR}(j)$ denotes the estimated impulse response function at horizon j to a unit innovation at time 0. Confidence intervals are computed using a non-parametric bootstrap procedure with 500 replications.[19] The half-life is appealing in that it has immediate intuition. But for completeness (and to help ensure robustness), we also report the largest autoregressive root in the processes we estimate (LAR), as well as the cumulated impulse response (CIR), which measures the total cumulative effect of a

[18] The only difference with the univariate case pertains to the innovation variances, which are here allowed to be country specific, so that $\sigma^2_{\tilde{q}_{ic}} = \sigma^2_{ic}/\left(1-\rho_i^2\right)$ and $\sigma_{\tilde{q}_{ic},\tilde{q}_{jc}} = \sigma_{ijc}/\left(1-\rho_i\rho_j\right)$.

[19] For the Arellano-Bond estimator, the bootstrap was performed using the Brown and Newey [2002] method.

unit shock to relative prices.[20] In all cases, bootstrapping procedures are implemented to derive standard error bands around our persistence estimates.

Below, we first present the panel estimators usually implemented in the real exchange rate literature, namely the fixed effects, Anderson-Hsiao and Arellano-Bond estimators. We show in section V that they reproduce standard results in our data. However, Section II showed none of these estimators is appropriate when there is sectoral heterogeneity in the dynamic parameters. So we next present two models allowing for heterogeneity, the Mean Group (MG) and the Random Coefficient (RC) estimators.

III.A. Standard Panel Estimators

We study both panels of disaggregated relative prices and panels of real exchange rates. We specify the latter as follows:

$$q_{ct} = \gamma_c + \sum_{k=1}^{K} \rho_k \; q_{ct-k} + \varepsilon_{ct}. \tag{4}$$

The possible presence of fixed effects through γ_c in equation (4) requires that the specification be estimated in first- or mean-differences. Moreover, the presence of a lagged dependent variable makes it necessary to use instrumental variables when estimating equation (4). Anderson and Hsiao [1982] proposed to instrument the differenced lagged dependent variable with its lagged level to alleviate the bias. The resulting instrumentation is often weak, which is why Arellano and Bond [1991] introduced a GMM procedure using all available lags as instruments of the differenced lagged dependent variable.

Correspondingly, we use a standard specification to investigate the speed of mean reversion in sectoral real exchange rates:

$$q_{ict} = \gamma_{ic} + \sum_{k=1}^{K} \rho_k \; q_{ict-k} + \varepsilon_{ict}, \tag{5}$$

which we estimate allowing for (generalized) fixed effects.

III.B. Heterogeneous Models

Our next step is to explore the cross-sectional heterogeneity of our panel. We allow for the possibility that

$$q_{ict} = \gamma_{ic} + \sum_{k=1}^{K} \rho_{ick} \; q_{ict-k} + e_{ict}, \tag{6}$$

where slopes and intercepts are allowed to vary across the panel units.

Both the Mean Group (MG) and the Random Coefficient (RC) models allow for heterogeneous coefficients. They differ in their assumptions on the nature of heterogeneity. While

[20] See Andrews [1993] for a discussion.

MG is a generalized Fixed Effects estimator that assumes 'deterministic' heterogeneity, RC is a generalized Random Effects estimator that allows individual specific random components in all estimated coefficients. In particular, it assumes $\gamma_{ic} = \gamma + \eta_{ic}^1$ and $\rho_{ick} = \rho_k + \eta_{ic}^2$, where η^1 and η^2 are assumed to have zero means and constant covariances. The RC model entails a Generalized Least Squares (GLS) procedure that optimally accounts for the stochastic heterogeneous nature of the residuals. In particular, GLS uses the variance-covariance matrix for η^1 and η^2 to weigh optimally the individual sector-specific slopes when aggregating them. The MG model introduced in Pesaran and Smith [1995] instead simply performs an arithmetic average of sector-specific slopes, with equal weights. The MG estimator is efficient if the optimal weights happen to be insignificantly different from the arithmetic ones. However, asymptotically, the two estimators are equivalent, as shown in Pesaran [2003]. This suggests a test procedure choosing between the two estimators, akin to the Hausman test used for standard panel estimates. In section V, we perform the relevant tests and let the data decide which estimator is the most appropriate.

The standard MG model estimates $\{\rho_k\}_{k=1}^K$, the mean autoregressive coefficients, by a simple arithmetic average of the least squares estimates for sector specific coefficients. We next describe the more general RC model, since MG is but a special case where heterogeneity is deterministic. Rewrite equation (6) as

$$q_{ict} = \gamma + \sum_{k=1}^{K} \rho_k \; q_{ict-k} + \varepsilon_{ict},$$

with $\varepsilon_{ict} = e_{ict} + \eta_{ic}^1 + \sum_{k=1}^K \eta_{ick}^2 q_{ict-k}$. Consistent GLS estimates of the coefficients of interest in equation (6) are given by an optimally weighted average of sector-specific point estimates. The analogy with the Random Effects estimator can best be seen by rewriting the model as

$$Q_{st} = Q_{sK} \; B_s + e_s,$$

where $Q_{st} = [q_{11t}, ..., q_{NCt}]'$, $Q_{sK} = [1, Q_{st-1}, ..., Q_{st-K}]$, $B_s = B + \eta_s$, with $B = [\gamma, \rho]'$, $\eta_s = [(\eta_s^1)', (\eta_s^2)']'$, and $\sigma_s^2 = E(e_s' \; e_s)$.[21] We have assumed that $E(\eta_s) = 0$. Further define $E(\eta_s \; \eta_s') = \Gamma$. The random coefficient estimator of B is given by

$$\begin{aligned} \widehat{B} &= \sum_s W_s \; B_{OLS}^s, \; W_s = \left[\sum_s (\Gamma + V_s)^{-1} \right]^{-1} (\Gamma + V_s)^{-1} \\ V_s &= \sigma_s^2 \left(Q_{sK}' \; Q_{sK} \right)^{-1}, \end{aligned}$$

where B_{OLS}^s denotes the sector specific OLS estimates of the slopes. Thus, RC applies the information in Γ efficiently when averaging the sector specific slopes.

In Section II, we allowed for non-zero cross-sectoral correlations in the residuals, with $E(\varepsilon_{it} \; \varepsilon_{jt}) = \sigma_{ij}$ for $i \neq j$. We shall want our estimator to allow for this possibility, too, since our proof showed that the magnitude of the bias depends on σ_{ij}. It is standard to implement the Seemingly Unrelated Regression (SURE-GLS) remedy to correct for cross-sectional correlations in error terms.

[21] This follows Hildreth and Houck [1968] and Swamy [1970, 1971].

It can be applied to either one of our heterogeneous estimators, but, since it estimates the covariance matrix of the residuals in the panel, SURE requires that the cross-sectional dimension be smaller than the time dimension of the data. This is unfortunately not the case in our data, where $N = 204$ and $T = 180$ in the full sample. We therefore need to truncate our data. We do so by using Engel's version of the Eurostat dataset, which has fewer observations than ours.[22]

Alternatively, Pesaran [2002] introduces a common correlated effects (CCE) estimator, well-tailored for large panels with both cross-sectional interdependence and heterogeneity. The estimator provides a correction to the MG estimator that accounts for unobserved common factors potentially correlated with individual-specific regressors. CCE allows for common effects in the residuals that can have a different impact on individual units, and that can be arbitrarily correlated amongst themselves. It is likely to improve on the SURE approach, as the estimation of the covariance matrix of the residuals has lower dimensionality, thanks to the structure imposed through the common effects. In CCE, we can include all cross-sections in our data, while keeping identification parsimonious. CCE may also yield more accurate estimates than SURE, since the latter is unable to capture, for instance, the effects of a persistent common factor on the residuals covariance matrix. Furthermore, the CCE estimator is straightforward to implement since the common effects correction of the MG estimator, for instance, simply amounts to including lagged cross-sectional averages in the least squares regressions performed by MG. In particular, the MG-CCE estimator determines ρ on the basis of the following regressions

$$q_{ict} = \gamma_{ic} + \sum_{k=1}^{K} \rho_{ick}\ q_{ict-k} + \sum_{h=0}^{H} \phi_{ich}\ \overline{q}_{t-h} + e_{ict},$$

where $\overline{q}$ is the cross-sectional average of q_{ic}. As for the standard MG estimator, the MG-CCE estimate of ρ_k is given by $\widehat{\rho}_k^{MG-CCE} = \sum_{ic=1}^{N} \widehat{\rho}_{ick}/N$. The cross-sectional averages $\overline{q}$ control for common shocks in the errors.

IV. Data

In this section, we first describe our data, including a discussion of their accuracy and representativeness. We then check that they verify the basic conditions for the bias discussed in Section II to be positive.

IV.A. Description

We study sectoral real exchange rates obtained from Eurostat, the statistical agency of the European Union. We focus on (non-harmonized) price indices for consumption goods and services, since harmonized price indices are available for very short samples only. The data correspond to monthly observations and cover at most the period 1960:1 to 2000:12. However, many observations are missing in the early and late part of the period, so we choose to focus on a [1981,1995]

[22] See www.ssc.wisc.edu/~cengel.

sample. This leaves us with a maximum of 180 time series observations (see Appendix 3). We report results based on checking the data in painstaking detail. In particular, we correct for obvious repetitions or outliers, and, whenever possible, use primary data sources obtained from national statistical agencies to correct the Eurostat data. Our modifications include those reported by Engel [2000b], but in some cases we were able to obtain the original data sources.[23] For completeness we also computed the results of our estimations based on Charles Engel's version of the Eurostat dataset, which we report in the text as well.

Eurostat publishes two-digit sectoral price indices for nineteen goods categories and thirteen countries. The goods categories are a mixture of low and high unit cost goods (e.g. bread and cereals versus vehicles), highly tradeable goods (e.g. clothing), goods commonly construed as non-tradeable in nature (public transport or hotels), and goods for which there is wide variation in the degree of product differentiation (fuel versus sound and photographic equipment).[24] Our sample thus constitutes an interesting cross-section with some variation along the dimensions commonly advanced to explain variations in relative prices. The cross-sectional variation is key to our analysis, since it allows us to identify the heterogeneity in relative price dynamics.

Our real exchange rates are CPI-based and defined against the US dollar.[25] Since our purpose is to investigate the effects of heterogeneity and aggregation, our sample of countries and the time coverage are identical for the two levels of aggregation. Furthermore, our measure of real exchange rates is based on the aggregation of the same exact sample of goods for which we have disaggregated information.[26] In particular, sectoral real exchange rates write

$$q_{ict} = \ln\left(\frac{S_{ct}\ P_{ict}}{P_{i,US,t}}\right),$$

where S_{ct} denotes the nominal bilateral exchange rate between country c and the US dollar at date t, P_{ict} is the price of good i in country c at date t, and $P_{i,US,t}$ is the corresponding US price. Aggregate real exchange rates, in turn, write

$$q_{ct} = \ln\left(\frac{S_{ct}\ P_{ct}}{P_{US,t}}\right).$$

We test for unit roots for both aggregate and sectoral real exchange rates. We use two panel data tests: Levin and Lin [1993] (henceforth LL) and Im, Pesaran and Shin [1997] (henceforth (IPS). The LL procedure tests the hypothesis that *all* the cross-sectional units are stationary against the hypothesis that they are *all* non-stationary. The IPS procedure is more general in

[23]For instance, we have additional observations for Finland and Greece. The data can be downloaded at http://faculty.london.edu/jimbs, http://www.princeton.edu/~hrey or http://faculty.london.edu/mravn.

[24]See Appendix 3 for details.

[25]All our conclusions stand if we use the British Pound as the anchor currency.

[26]We also checked CPI based real exchange rate measures based on the International Financial Statistics database released by the IMF. The results were almost identical.

that it allows for *some*, but not all, of the series to be stationary under the alternative hypothesis. Table I reports the results for the two tests. The first column reports the outcome of several unit root tests for the panel of aggregate real exchange rates. The third column concerns the panel of sectoral real exchange rates, and, for completeness, the second column pertains to the nominal exchange rate. In each case we report the IPS test and two variations of the LL test, allowing or not for individual effects. Each estimation is performed both with or without a trend term. Table I shows that the evidence tilts in favor of stationarity in both the real exchange rates and sectoral relative prices, with four out of five test statistics supporting stationarity in each case.[27] This is consistent with the findings of an enormous literature.[28]

IV.B Conditions for a Positive Bias

In sections II.A and II.B, we derived conditions for the positivity of the bias. Whenever the covariance between the coefficients α_i or δ_i and the persistence parameters ρ_i is positive, the bias is positive, since more persistent sectors get higher weights on average in the estimation. We now verify whether these conditions hold in our data.

For each sector in each country of our sample, we retrieve estimates of the autoregressive parameter, as well as the variance covariance matrix of the sectoral innovations. We do this by estimating AR(1) processes for each panel unit separately:

$$q_{ict} = \gamma_{ic} + \rho_{ic}\ q_{ict-1} + e_{ict}.$$

We use these estimates to compute the coefficients α_i and δ_i on the basis of the formulas derived in sections II.A and II.B. We stress these results are simply first checks as to whether the bias discussed in Section II is positive or not. Our final estimates are contained in Section V.

Figure I plots the coefficients α_i and δ_i against estimates of the autoregressive parameters ρ_i.[29] In both cases we find a strong positive correlation between the weights and the persistence parameters. The covariance between α_i and ρ_i is equal to 0.166. It is equal to 0.229 for δ_i.[30] Furthermore, Δ is in all cases unambiguously positive when computed on the basis of these estimates. Thus, we expect a positive bias in persistence estimates arising from panel estimators (including aggregate ones) that do not allow for dynamic heterogeneity.

---

[27] The IPS test fails to reject non-stationarity of disaggregate relative prices when a trend is included, and of the real exchange rate when no trend is included. But standard tests also reject the presence of a trend in relative prices, and suggest there may be one in the real exchange rate.

[28] Frankel and Rose [1996], Oh [1996], Wu [1996] or Lothian [1997] all reject non-stationarity in a variety of cross-country panel datasets. Using Monte-Carlo evidence, Engel [2000a] argues standard tests may be unable to detect unit roots in real exchange rates in the presence of a stationary, but noisy, component. But Ng and Perron [2002] take this into account and estimate a half-life for real exchange rate shocks between nine and fifteen quarters, right back in the consensus view.

[29] We do not pool across sectors. Therefore, Figure I contains as many datapoints as the cross-sectional dimension of our dataset.

[30] The weights used are the Eurostat harmonized indices of consumer prices weights. When equal weights are used the covariance is 0.223.

V. Aggregation Bias in Practice: PPP Strikes Back

This section investigates empirically the importance of heterogeneity by comparing results derived from standard methods to those obtained from estimators allowing for dynamic heterogeneity. We first review aggregate results and confirm that our data are not particular in any way, as we are able to reproduce consensus estimates. We then implement heterogeneous estimators and find substantially faster mean reversion in relative prices.

V.A. Results for panels of Aggregate Real Exchange Rates

We first estimate equation (4) using real exchange rates vis-a-vis the US dollar. This corresponds directly to standard estimates of real exchange rate persistence based on panels of real exchange rates. The results are reported in Table II. Lag lengths were identified using a general-to-specific technique starting from a maximum of twenty lags.[31] We report two tests for parameter homogeneity: a Hausman-type test and that proposed by Swamy [1971]. Neither of them can reject the null hypothesis that the dynamics of the panel units are homogeneous (across countries) at conventional levels of confidence, a finding that is consistent with the evidence in Boyd and Smith [1999]. We stress that this does not preclude heterogeneity in panels of *sectoral* real exchange rates.

The first row in Table II reports the results based on the OLS fixed effects panel estimator. The estimates imply a half-life roughly at the center of the consensus view, with a point estimate of three years and ten months. The bootstrapped 95 percent confidence interval ranges from around two and a half years to just below five years. The alternative measures we report also imply high persistence. The largest autoregressive root for example has a point estimate of 0.97. These results are entirely in line with existing results. The Table also reports that the presence of fixed effects cannot be rejected, and a Hausman test favors the fixed effects specification over a random effects model.

The presence of fixed effects demands that the model be estimated in first differences. As mentioned earlier, due to the presence of lagged dependent variables this produces correlation between errors and the regressors, which requires instrumenting the lagged dependent variable. Both the Anderson-Hsiao IV-type estimator and the Arellano-Bond GMM estimator lead to a significant upward revision in the estimate of real exchange rate persistence. As far as the Anderson-Hsiao estimator is concerned, the implied half-life is six years. The 95 percent confidence interval ranges from just below three years to infinity, but this is probably due to poor small-sample properties (and a large root mean square error).[32] The Arellano-Bond estimator on the other hand, has both better small-sample properties and a lower root mean square error. Estimates imply a half-life of four and a half years, with a 95 percent confidence interval between just below four years and just above six years. Our aggregate results are in agreement with the

[31] Twelve lags in the case of the Anderson-Hsiao estimator.

[32] Lagged relative prices make for weak instruments, and hence a poor first-stage fit.

existing literature. The aggregate dimension of our data generates perfectly standard results. This is reassuring for our data pertain to European exchange rates vis-a-vis the United States (or the United Kingdom), all developed economies for which integration could result in faster reversion to price parity in general.

V.B. Results for Sectoral Real Exchange Rates

We now investigate the results based on the panel of disaggregated prices. We work with exactly the same panel of sectoral prices that compose the aggregate CPI used in the aggregate analysis in the previous section. We use six alternative estimates. First, simple fixed effects estimates, which would be valid under homogeneity. We then extend the fixed effects estimator to allow for cross-sectional dependence. We use either a SURE approach, or the adjustment for common effects introduced by Pesaran [2002]. Second, we check which heterogeneous estimator is applicable to our data, RC or MG. We then present results for the same three variations of the preferred heterogeneous estimator.

Table III summarizes all results. The fixed effects estimator, which does not allow for sectoral dynamic heterogeneity, implies a half-life of three years. The estimates are relatively precise, as the 95 percent confidence interval ranges from roughly two and a half to three and a half years. This lies at the lower end of the consensus view but does not differ markedly from previous results in the literature. We implement both the SURE and the CCE corrections to the fixed effects estimator, but the estimates remain largely unaffected. This suggests correction for correlated residuals does not bring down persistence estimates based on homogeneous estimators.

But both the Hausman and Swamy tests indicate clear rejection of the hypothesis of homogeneity of the slope coefficients across sectors, at any level of confidence.[33] This immediately implies that the fixed effects estimator is inconsistent, as discussed in section II. A more subtle implication concerns the results reported in Table II, which *also* suffer from the 'aggregation bias', as showed in Section II, even though there is no heterogeneity across countries for aggregate exchange rate dynamics. This happens since standard panel estimators assign larger weights to the components of the real exchange rate that display slower mean reversion. It still remains to be seen, however, whether the bias is important quantitatively.

To that end, we now turn to estimators designed to account for dynamic heterogeneity. Table III reports the test introduced in Pudney [1978] meant to assess if the data support a Random Coefficient or a Mean Group model. As is clear from the Table, the data resoundingly reject the Random Coefficient model. In what follows we therefore use the Mean Group (MG) estimator.

The MG estimator produces a half-life just above two years, and a 95 percent confidence interval ranging from fourteen to twenty eight months. This is already significantly below the 'consensus view', an interesting outcome in itself for it suggests the aggregation bias is large and prevalent in our data.[34] Further, the 95 percent confidence interval for the largest autoregressive

[33] The Hausman test allows for correlated residuals, and is based on the CCE correction. An alternative based on the SURE correction has value 57.68, with a P-value equal to zero.

[34] To compute confidence intervals for our heterogeneous estimators, we use the mean coefficients to draw the

root now ranges from 0.903 to 0.973. The MG estimates are almost all significantly distinct from the intervals obtained using (homogeneous) fixed effects, whether on aggregated or on disaggregated data. This is true for both half-life measures and estimates of the largest root.

This leaves open the question of correlated residuals. Table III reports a Breusch-Pagan test checking the diagonality of the covariance matrix of the residuals, as implied by the MG regressions. The null-hypothesis of diagonality is overwhelmingly rejected. As mentioned in Section II, correlated residuals can affect the magnitude of the bias. A common prior is that price movements tend to synchronize across sectors, which suggest our corrected MG estimates should yield even lower measures of persistence.[35] This is confirmed by both our SURE and CCE estimates.[36]

Allowing for cross-sectional dependence through the use of the MG SURE estimator lowers the point estimate of the half-life of relative prices to below two years (22 months), estimated with precision since our confidence interval ranges from 17 to 27 months. This is significantly below the consensus view.[37] However, implementing the SURE estimate requires (arbitrary) truncation of our dataset, which otherwise contains too many cross-sections for its time dimension. On the other hand, the MG CCE estimator can be implemented on our preferred dataset, and it implies a half-life point estimate of eleven months. The 95 percent confidence interval ranges from seven to twelve months. According to this estimator, an upper bound for the real exchange rate half-life is one year.

Given the importance of heterogeneity to our results, we dedicate a few lines to describing our sectoral persistence estimates.[38] Figure II plots the distribution of sector-specific persistence as measured by the largest auto-regressive root (LAR) implied by our MG CCE estimates. For each sector, we report the mean root (denoted with a diamond sign) and the interquartile range of the sectoral estimates across countries. Heterogeneity occurs at two levels. First, there is substantial heterogeneity of persistence within countries, with large differences in mean LAR across goods, as confirmed by the British example, indicated in the Figure by a cross. Second, some specific sectoral exchange rates also display heterogeneous dynamics across countries. The latter aspect makes the panel approach attractive.

Food products (such as bread, meat, dairy or alcohol), as well as domestic appliances, fuel or furniture tend to display low persistence on average across countries. On the other hand,

---

residuals, and then perform sampling from the residuals themselves. This was suggested to us by Ron Smith.

[35] Actually, in our data, there is not a single instance of non-positive covariances between sectoral price residuals.

[36] We also confirm that all our results hold in Charles Engel's version of the Eurostat data set: the MG gives a half life of 25 months (confidence interval 9-31); the MG CCE a half life of 13 (confidence interval 9-24).

[37] Our SURE estimates are based on a sample where N is only marginally smaller than T. Given this dimensionality, that we should find such low and precise estimates suggests common effects are strongly present in our data.

[38] A caveat is in order for what follows. The precision of our *aggregate* estimates (MG, RC and otherwise) is mostly afforded thanks to our large cross-section. Sector-specific estimates are substantially less efficient and precise, and should be used for illustrative purposes rather than for inference.

persistence tends to be higher in clothing, footwear, vehicles and rents. Despite substantial dispersion across countries, our estimates square overall relatively well with a heuristic classification of sectors in traded and non-traded activities. For instance, relative prices in rents, hotels, or vehicles are on average persistent, whereas the relative prices of bread, meat, dairy, tobacco, fuel or alcohol are not.

Does Figure II mirror existing studies of disaggregated relative prices? First and foremost, it is important to note that one recurrent conclusion in most of the existing work is heterogeneity, both across sectors and across countries (for sectoral exchange rates). Like us, Cheung et al. [2001] find evidence of substantial heterogeneity in persistence for their panel of fourteen OECD country sectoral real exchange rates. Knetter [1993] documents heterogeneity in exchange rate pass-through across goods and export market destinations. He also finds low pass-through of exchange rates into car prices, and is surprised to find high pass-through in alcohol; these findings are consistent with our results. Like us, Yang [1997] finds substantial heterogeneity across sectors, with low pass-through in apparel and other textiles, and high pass-through in capital goods. These results are consistent with our estimates of persistence for clothing, footwear and domestic appliances. Campa and Goldberg [2002] find a relatively high degree of pass-through for food items, which is consistent with our low persistence estimates for dairy, bread, meat and fruit. Similarly, Crucini and Shintani [2002] estimate good-by-good persistence for a sample of cities in different countries. They find overall rapid price adjustment. Their selection of goods, drawn from the Economist Intelligence Unit, is dominated by food and beverages, household services and furnishings. These are exactly the type of goods for which we also tend to find low mean persistence and indeed overall relatively low cross-country dispersion.

V.C. Monte Carlo Experiments

This section explores robustness along two dimensions. First, we compare the abilities of various estimators to capture the heterogeneity bias. This is particularly relevant, for it enables us to compare standard panels to heterogeneous estimators on the one hand, and standard MG to its SURE and CCE variants on the other. Second, we let the extent of heterogeneity and persistence vary between plausible bounds and ask how the magnitude of the bias responds.

In Figure III, we ask how the standard panel estimators would have performed were the data generating process the one we estimated using the MG-CCE. Our focus is on the bias affecting the first autoregressive parameter. It is clear that the MG CCE estimator has satisfactory properties. The only other approach that appears to be consistent is the MG SURE estimator. All others induce a large positive bias. In the cases of Fixed Effects and the Anderson-Hsiao variants, this can be traced back to the failure to account for dynamic heterogeneity. Allowing for cross-sectional error correlation does not improve the FE estimator's properties. Thus, homogeneous estimators induce a bias in persistence estimates in the presence of heterogeneity in the dynamics of the panel units - whether correlated residuals are accounted for or not. The uncorrected MG estimator also appears positively biased, though marginally less than FE. This bias stems from positively correlated residuals, i.e. positive realizations of σ_{ij} in Δ. Finally, the Figure shows the presence of a large positive bias in the aggregate estimator as well, confirming that *sectoral* heterogeneity translates into an *aggregate* bias.

In Figure IV, we use a wide range of alternative data generating processes to illustrate the magnitude of the aggregation bias along two dimensions: the underlying persistence of the data generating process, and the underlying heterogeneity. As predicted by theory, the aggregation bias increases with the extent of heterogeneity, as measured by the cross-sectional variance of η_i, irrespective of the estimator implemented. Further, the aggregation bias continues to dominate even at high levels of persistence.[39] We come back to this point at length in Section VI.

V.D. PPP Puzzle

The empirical importance of heterogeneity has two substantive consequences. First, from a theoretical standpoint, models with built-in heterogeneity ought to be able to generate endogenous persistence, for instance in the real exchange rate. This possibility has recently been explored by several authors.[40] Second, from an empirical standpoint, estimates that control for the heterogeneity present in the data ought to be closer in magnitude to what is implied by one-sector models. We conduct a very simple experiment. Using the MG CCE estimates as a data generating process, we compare the persistence implied by our estimates to the simulation results obtained in a recent one-sector two-country model, due to Chari, Kehoe and McGrattan [2002]. Their baseline calibration of a sticky price model implies a first order autocorrelation of 0.62, while our estimated model implies a coefficient of 0.60, with a standard deviation of 0.07. Our corrected estimates appear consistent with plausible nominal rigidities.[41] Thus, we arguably solve the PPP puzzle, at least in our data set.

Recent contributions have studied the impact of non-linearities on the estimation of half-lives.[42] We have restricted ourselves to the effects of heterogeneity in persistence in the context of linear autoregressive models. Undoubtedly, it would be of interest to extend our analysis to non-linear settings. However, very little is known about the effects of aggregation, let alone heterogeneity in the presence of non-linearities. But our results are closely related to the literature on non-linearities in aggregate real exchange rates. As discussed earlier, heterogeneous dynamics may give rise to long memory in aggregate real exchange rates. Diebold and Inoue [2001] show that the dynamics produced by long memory models may be arbitrarily close to those produced by non-linear models.[43] Non-linear dynamics of *aggregate* real exchange rates may be fully compatible with -or at least observationally equivalent to- our argument about the importance

[39] None of the stochastic processes entering our simulations have autoregressive coefficients above one.

[40] For instance Ghironi and Melitz [2004], Bergin and Glick [2003], Lewbel and Ng [2003] or Ravn [2001].

[41] We generated 1000 time series each with 160 quarterly observations (by point-in-time sampling the monthly data). Since Chari et al. [2002] HP-filter their data, we implemented the filter on our simulated data as well.

[42] For instance, Obstfeld and Taylor [1997], Michael, Nobay and Peel [1997], Taylor and Peel [2000], Taylor, Peel and Sarno [2001], Kilian and Taylor [2003] and Imbs, Mumtaz, Ravn and Rey [2003].

[43] Diebold and Inoue [2001] discuss the relationship between long memory and non-linear models such as mixture, permanent break, and Markov-switching models. They conjecture that a similar relationship may exist with threshold autoregressive models.

of heterogeneity at the disaggregated level.

VI. Robustness Checks

In this section we investigate the robustness of our findings. We first assess the importance of measurement error in sectoral data. We then evaluate the importance of another (attenuating) bias, recently emphasized in the empirical exchange rate literature.[44]

VI.A. Errors in Variables

There is a presumption that measurement error is more prevalent in sectoral data than in the aggregate. Indeed, if errors are uncorrelated across sectors, they tend to average away in the aggregate, and the resulting attenuating bias that may arise from examining disaggregated data might explain the discrepancy we just documented. However, as we illustrated in Section IV.B and confirmed in Figure II, we do not observe *systematically* low half-lives at the sectoral level, and this casts doubt on the alternative explanation right at the outset. We do however also address the issue in a classic econometric manner. In the absence of measurement error, the OLS estimator of persistence, ρ^{OLS} and an instrumental variable estimator ρ^{TSLS} are both consistent, and the OLS estimator is efficient. However, in the presence of measurement error, the OLS estimator is inconsistent. Therefore, $plim\ (\rho^{OLS} - \rho^{TSLS})$ should be non zero in the presence of measurement errors. We perform a Hausman test along those lines, but take into account parameter heterogeneity. In particular, we carry out these tests at the sectoral level for each of the cross-sectional units.

Let q^*_{it} denote the observed value of the sectoral real exchange rate and q_{it} its true value. u_{it} denotes measurement error. The model is given by

$$q^*_{ict} = \gamma_{ic} + \sum_{k=1}^{K} \rho_{ik}\ q^*_{ict-k} + \nu_{ict}$$

where $q^*_{ict} = q_{ict} + u_{ict}$ and $\nu_{ict} = -u_{ict} + \sum_{k=1}^{K} \rho_{ik}\ u_{ict-k} + \varepsilon_{ict}$. The lag structure of the model implies that $\{q_{ict}, .., q_{ict-K}\}$ are correlated with the error term $\nu_{ict.}$ Appropriate instruments for the TSLS estimate are therefore $\{q_{it-K-1}, ..., q_{it-2K}\}$. Hausman tests (available upon request) indicate the null hypothesis that OLS is consistent is rejected for only one panel unit. This makes it doubtful that an errors-in-variables bias is relevant in our data.

[44] We checked robustness along several other dimensions but do not report all results. For instance, we used GBP as our anchor currency. Our Mean Group based half-life point estimate was then seventeen months, down to fourteen months when corrected using CCE. We also verified that our results obtain even if we constrain the lag length to be the same across estimators. For instance, with *one* lag, the FE half-life equals 36 months, MG's is 27 months and MG-CCE's is 16 months.

VI.B. Downward Bias in OLS for Highly Persistent Processes

For highly persistent autoregressive processes, it is well known that least squares estimators may be biased downward in small samples.[45] This bias could counter our claim that the PPP puzzle is due to a heterogeneity bias. Furthermore, the least squares bias may persist even as the cross-sectional dimension of our panel, N, rises to infinity (Nickell [1981]). This possibility has recently received considerable attention in the PPP literature. Murray and Papell [2002a] apply an approximately median bias-corrected estimator proposed by Andrews and Chen [1994] to study the persistence of real exchange rates. Their estimates of the (approximately median unbiased) confidence intervals are so wide that they conclude the data are basically uninformative about the half-life of the real exchange rate.[46] Murray and Papell [2002b] apply similar methods to a panel of real exchange rates. Their bias reduced half-lives estimates lie exactly in the range of three to five years, right back in the consensus view.[47] Finally Murray and Papell [2002c] apply again the Andrews-Chen correction to a single exchange rate series (dollar-sterling), and argue that previous results were misguided. Their 95 percent confidence intervals for the half-life range between three to five years and infinity. Rossi [2003] instead applies local-to-unity asymptotic theory to construct confidence intervals for the half life. Interestingly, although the upper bounds on the bias reduced real exchange rate persistence are still high, she finds that the lower bounds are very low, and her confidence interval does not exclude the consensus view.

The attenuating bias might be important in the present context, as it could contribute to explaining our surprisingly low estimates. Very little is known, however, about the joint effects of the heterogeneity bias we highlight and the small sample bias stressed in the literature. In an interesting paper, Choi, Mark and Sul [2003] evaluate the relative importance of these two biases in the context of simulations.[48] They implement a Monte Carlo experiment with heterogeneous dynamics, and find a tendency for the overall bias in fixed effects estimators to be negative. There are three reasons why our estimates do not fall directly victim to these *simulation* results. First, their data generating processes do not allow for the common correlated effects that we find are important in our data, and pertinent for our results. Second, the dominance of small sample bias appears sensitive to the parametrization of heterogeneity and the length of the artificial data. Third, the results of the Monte Carlo simulations are sensitive to initial conditions.[49] As far as

[45] See for instance Hurwicz [1950] or Orcutt [1948].

[46] Qualitatively similar results have also been obtained by Kilian and Zha [2002] using a different methodology. See also Elliott and Stock [2001].

[47] This is also the conclusion in Cashin and McDermott [2003], who allow for a moving average error structure. They show that real exchange rates half-lives remain firmly - and significantly - within the consensus range, even after correcting for the bias.

[48] Most of their analysis is centered around evaluating the relative importance of the small sample and temporal aggregation biases in actual data. Homogeneity cannot be rejected in their data, and they use pooled estimators to implement their bias correction method. Thus, the only conclusion they draw from actual data is that the small sample bias dominates the temporal aggregation bias.

[49] See Imbs, Mumtaz, Ravn and Rey [2004] for more details.

estimations are concerned, Choi, Mark and Sul [2003] do not implement their correction methods directly on *heterogeneous* data. They carefully test for heterogeneity in their panel of aggregate exchange rates and fail to reject homogeneity (as we do for aggregate real exchange rates). In contrast, we have to implement bias correction techniques in the presence of heterogeneous dynamics in our panel of *sectoral* real exchange rates.

In Imbs, Mumtaz, Ravn and Rey [2004] we examine the properties of various bias reduction techniques on the basis of Monte Carlo experiments. Importantly, we do allow for common correlated components in the data generating process. We show that ignoring these effects actually results in a serious *positive* bias in *corrected* half-life estimates based on the MG estimator.[50]

Our simulations apply two versions of the Kilian [1998] 'bootstrap-after-bootstrap' procedure to correct the half-life estimated on the basis of the MG estimator (as well as its SURE and CCE refinements). We first calculate a bias-corrected estimate of the half-life on the basis of the corrected autoregressive coefficients, an *indirect* approach followed by Chen and Engel [2004]. Alternatively, we correct the estimated half-life *directly* using the bootstrap algorithm. We find that the latter method outperforms the former, a result related to Pesaran and Zhao [1999] who argue the correction should be directly applied to the object of interest in heterogeneous dynamic panels.[51]

In Table IV, we report half life estimates that were corrected for small sample bias using Kilian's method on our panel of sectoral real exchange rates.[52] We implemented both the direct and indirect approaches. For the standard MG estimates, the corrected half-life is higher and has wider confidence intervals, especially if computed using the indirect approach. This is entirely consistent with the Monte Carlo experiments in Imbs, Mumtaz, Ravn and Rey [2004], who show that failing to control for common correlated effects biases upwards MG estimates corrected for small sample bias. When we turn to the proper estimator, the MG CCE model, our corrected estimates point to a half-life of eighteen months. This constitutes a marginal increase relative to our original (uncorrected) estimate of eleven months.[53] The 95 percent confidence interval is narrow - spanning eleven to twenty-eight months - and still excludes the 'consensus view'. Small sample bias is potentially a relevant concern in our context, but our results seem to remain largely unchanged when our estimates are corrected. In our data, the dynamic heterogeneity bias dominates.

[50] The same is true for the So and Shin [1999] method.

[51] For more details see Imbs, Mumtaz, Ravn and Rey [2004].

[52] Phillips and Sul [2003] propose an alternative bias reduction method, the Panel Feasible Generalized Median Unbiased estimator applicable to heterogeneous panels. This method relies upon applying a median unbiased correction to a SUR panel estimator. Given the large cross-sectional dimension of our data, we apply instead the 'bootstrap-after-bootstrap' procedure.

[53] This result is based on the 'direct approach'. As far as the MG CCE estimator is concerned, the 'indirect approach' gives similar results, with a corrected half-life estimate of twenty months. The MG-SURE gives a half life of twenty-seven months when the direct approach is used. We show in Imbs, Mumtaz, Ravn and Rey [2004] that this estimate may also be biased upwards.

VII. Summary and Conclusions

We have argued a simple mechanism may explain the difficulty in reconciling real exchange rate dynamics with the predictions of models with realistic impediments to price adjustment. Our argument rests on the possibility that relative price dynamics differ across the goods composing the real exchange rate. If for instance goods differ substantially in their tradability, the degree of competition or transportation costs, there is little reason to expect a priori that relative prices converge homogeneously. Under this premise, the paper shows that the persistence of the real exchange rate should be interpreted as a biased estimate of the average persistence in relative prices. Under conditions which hold in the data, the bias is positive. Our results do not imply nor require that disaggregated relative prices all converge faster than the aggregate real exchange rate. If relative prices all converged quickly, so would the aggregate and there would be no bias, nor, indeed, any PPP puzzle. In reality some prices converge slowly and others do quickly. The "aggregation bias" comes precisely from this heterogeneity in dynamics.

Our data reproduce consensual estimates for real exchange rate persistence when standard panel techniques are implemented. They do as well when price dynamics are constrained to be identical across different goods. But this constraint is actually rejected in our data, and when accounting for heterogeneous dynamics, our measure of average persistence falls dramatically. Our estimates point to a half-life for sectoral real exchange rates down to between eleven and eighteen months, with tight confidence intervals that exclude the three to five year "consensus view" summarized in Rogoff [1996]. Such low estimates are consistent with realistic degrees of nominal rigidity. Thus, at least in our dataset, we appear to solve the PPP puzzle.

Our results seem robust. They withstand numerous alterations, truncations or variations to our dataset. They cannot be explained by the presence of measurement errors. They survive small sample bias corrections. Our corrected persistence estimates are only moderately larger than eleven months, up to eighteen months. And the confidence interval remains significantly below the "consensus view". We recognize that our sample is limited to European exchange rates vis-a-vis the United States (or the United Kingdom). Our focus is on developed and integrated economies. Whether our methodology will be useful for a broader set of countries, including in particular emerging markets, still remains to be seen.

Our findings have potentially important implications. First, our estimates for the average persistence in relative prices can be reproduced in models with realistic price rigidities. We bridge the gap between theory and evidence in this area of international macroeconomics and show persistence estimates based on disaggregated price data do not necessarily translate into similar results in the aggregate. Second, we underline the importance of heterogeneity at the microeconomic level for understanding macroeconomic aggregate phenomena. When microeconomic heterogeneity is purged from the data used to evaluate them, macroeconomic models perform better, at least as far as the real exchange rate is concerned. By the same token, whether models with non-trivial sectoral heterogeneity are capable of mimicking aggregate data is in our opinion an exciting area for future research. Should such models prove unsuccessful at generating persistent real exchange rates, there would indeed still be a PPP puzzle.

Appendix 1: Proofs

Proof of Corollary 1.1

We want to show that a sufficient condition for the dynamic heterogeneity bias to be positive is $0 \leq \alpha_i \leq \alpha_{i+1}$ for all i. These inequalities constitute a strong sufficient condition that ensures the positivity of Δ. Observe that $\sum_{i=1}^{N} \alpha_i = 1$ so that the α_i constitute *convex weights.* Our assumption on the ordering of the ρ_is implies that $0 \leq \alpha_i \leq \alpha_{i+1}$. Therefore we have

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

since two sets of *convex* combinations of the ρ_is are compared, one with increasing weights and one with equal weights. The left hand side convex combination, involving the α_is, gives higher weights to the largest ρ_is, while it gives equal weights $\frac{1}{N}$ to the ρ_is on the right hand side of the inequality. Hence the inequality holds. Then $1/N \sum_{i=1}^{N} \rho_i = \rho = \sum_{i=1}^{N} \rho \alpha_i$ implies

$$\sum_{i=1}^{N} \rho_i \alpha_i \geq \sum_{i=1}^{N} \rho \alpha_i,$$

which is equivalent to:

$$\sum_{i=1}^{N} \alpha_i \left(\rho_i - \rho\right) \geq 0.$$

Therefore, $\Delta = \sum_{i=1}^{N} \left(\rho_i - \rho\right) \alpha_i \geq 0$ and the bias is positive.

Proof of Corollary 1.2

This comes directly from the expression of the bias, $\Delta = \left(1/N \sum_{i=1}^{N} \left(\rho_i - \rho\right) \frac{\sigma_i^2}{1-\rho_i^2} \right) / \left(1/N \sum_{i=1}^{N} \left(\frac{\sigma_i^2}{1-\rho_i^2} \right) \right)$ whose magnitude increases on average with $(\rho_i - \rho)$, the distance between sectoral persistence to the cross-sectional average, holding $\{\sigma_i^2\}$ constant.

Appendix 2. Bias: Analytics

We first derive the (asymptotic properties for the) least squares estimate of the first-order autoregressive coefficient of q_t, given by $\rho^Q = E\left(q_t q_{t-1}\right)/E\left(q_t^2\right)$.[54] Let us first derive the expression of the bias assuming constant weights to economize on notations ($\omega_i = \omega_j$ for all i). We have:

$$\begin{aligned} E\left(q_t^2\right)_{(\omega_i=\omega_j)} &= \frac{1}{N^2}\left(\sum_{i=1}^{N}\sigma_{q_i}^2 + 2\sum_{i<j}^{N}\sigma_{q_i,q_j}\right) \\ E\left(q_t q_{t-1}\right)_{(\omega_i=\omega_j)} &= \frac{1}{N^2}\left(\sum_{i=1}^{N}\left(\rho_i\sigma_{q_i}^2 + \sum_{i<j}^{N}\left(\rho_i+\rho_j\right)\sigma_{q_i,q_j}\right)\right) \end{aligned}$$

$$\begin{aligned} \text{plim}_{N\to\infty,T\to\infty}\left(\rho^Q_{(\omega_i=\omega_j)}\right) &= \frac{\sum_{i=1}^{N}\rho_i\sigma_{q_i}^2 + \sum_{i<j}^{N}\left(\rho_i+\rho_j\right)\sigma_{q_i,q_j}}{\sum_{i=1}^{N}\left(\sigma_{q_i}^2 + 2\sum_{i<j}^{N}\sigma_{q_i,q_j}\right)} \\ &= \rho - \rho + \frac{\sum_{i=1}^{N}\rho_i\sigma_{q_i}^2 + \sum_{i<j}^{N}\left(\rho_i+\rho_j\right)\sigma_{q_i,q_j}}{\sum_{i=1}^{N}\left(\sigma_{q_i}^2 + 2\sum_{i<j}^{N}\sigma_{q_i,q_j}\right)} \\ &= \rho + \frac{\sum_{i=1}^{N}\left(\rho_i-\rho\right)\sigma_{q_i}^2 + \sum_{i<j}^{N}\left[\left(\rho_i-\rho\right)\sigma_{q_i,q_j} + \left(\rho_j-\rho\right)\sigma_{q_i,q_j}\right]}{\sum_{i=1}^{N}\left(\sigma_{q_i}^2 + 2\sum_{i<j}^{N}\sigma_{q_i,q_j}\right)} \\ &= \rho + \frac{\sum_{i=1}^{N}\frac{\rho_i-\rho}{1-\rho_i^2}\sigma_{\ i}^2 + \sum_{i<j}^{N}\left(\frac{\rho_i-\rho}{1-\rho_i\rho_j}\sigma_{\ ij} + \frac{\rho_j-\rho}{1-\rho_i\rho_j}\sigma_{\ ij}\right)}{\sum_{i=1}^{N}\left(\sigma_{q_i}^2 + 2\sum_{i<j}^{N}\sigma_{q_i,q_j}\right)} \end{aligned}$$

Reintroducing non equal weights (which amounts to multiplying variance by ω_i^2 and covariances by $\omega_i\omega_j$), it follows trivially that

$$\text{plim}_{N\to\infty,T\to\infty}\left(\rho^Q\right) = \rho + \frac{\sum_{i=1}^{N}\frac{(\rho_i-\rho)\omega_i^2}{1-\rho_i^2}\sigma_{\ i}^2 + \sum_{i<j}^{N}\left(\frac{(\rho_i-\rho)\omega_i\omega_j}{1-\rho_i\rho_j}\sigma_{\ ij} + \frac{(\rho_j-\rho)\omega_i\omega_j}{1-\rho_i\rho_j}\sigma_{\ ij}\right)}{\left(\sum_{i=1}^{N}\left(\omega_i^2\sigma_{q_i}^2 + 2\sum_{i<j}^{N}\omega_i\omega_j\sigma_{q_i,q_j}\right)\right)}$$

which we rewrite as[55]

$$\rho^Q = \rho + \Delta$$

with $\Delta = \left[\sum_{i=1}^{N}\frac{\omega_i^2}{1-\rho_i^2}\sigma_{\ i}^2 + \sum_{i\neq j}^{N}\frac{\omega_i\omega_j\sigma_{ij}}{1-\rho_i\rho_j}\right]\frac{(\rho_i-\rho)}{\sum_{i=1}^{N}\left(\frac{\omega_i^2}{1-\rho_i^2}\sigma_{\ i}^2 + \sum_{i\neq j}^{N}\frac{\omega_i\omega_j\sigma_{ij}}{1-\rho_i\rho_j}\right)}$.

[54] This derivation was first presented in Imbs, Mumtaz, Ravn and Rey [2003].

[55] This last expression is the same as the one presented in Chen and Engel [2004], who also generalize slightly Imbs, Mumtaz, Ravn and Rey [2003] by including weights in the derivation of the bias.

Appendix 3. Data Coverage

	BE	DE	DK	ES	IT	FR
Bread	81-1:95-12	81-1:95-7	81-1:95-12	81-1:95-12	81-1:95-12	81-1:95-12
Meat	81-1:95-12	81-1:95-7	81-1:95-12	81-1:95-12	81-1:95-12	81-1:95-12
Dairy	81-1:95-12	81-1:95-7	81-1:95-12	81-1:95-12	81-1:95-12	81-1:95-12
Fruits	81-1:95-12	81-1:95-7	81-1:95-12	81-1:95-12	81-1:95-12	81-1:95-12
Tobacco	81-1:95-12	81-1:95-7	81-1:95-12	81-1:95-12	81-1:9:1-12	81-1:95-12
Alcohol	81-1:95-12	81-1:95-7	81-1:95-12	81-1:95-12	81-1:95-12	81-1:95-12
Clothing	81-1:95-12	81-1:95-7	81-1:95-12	81-1:95-12	81-1:95-12	81-1:95-12
Footwear	81-1:95-12	81-1:95-7	81-1:95-12	81-1:95-12	81-1:95-12	81-1:95-12
Rents	81-1:95-12	81-1:95-7	81-1:95-12	81-1:95-12	81-1:95-12	81-1:95-12
Fuel	81-1:95-12	81-1:95-7	81-1:95-10	81-1:95-12	81-1:95-12	81-1:95-12
Furnit.	81-1:94-10	81-1:94-10	81-1:94-10	81-1:94-10	81-1:94-9	81-1:94-10
Dom. Appl.	81-1:94-10	81-1:94-10	81-1:94-10	81-1:94-10	81-1:94-9	81-1:94-10
Vehicles	81-1:95-12	81-1:95-7	81-1:95-12	81-1:95-12	81-1:95-12	81-1:95-12
Pub. Transp	81-1:95-12	81-1:95-7	81-1:95-12	81-1:95-12	81-1:95-12	81-1:95-12
Comm.	81-1:95-12	81-1:95-7	81-1:95-12	81-1:95-12	81-1:95-12	81-1:95-12
Sound	81-1:95-12	81-1:95-7	81-1:95-12	81-1:95-12	81-1:95-12	81-1:95-12
Leisure	81-1:95-9	81-1:95-7	81-1:95-9	81-1:95-9	81-1:95-9	81-1:95-9
Books	81-1:95-9	81-1:95-7	81-1:95-9	81-1:95-9	81-1:95-9	81-1:95-9
Hotels	81-1:95-9	81-1:95-7	81-1:95-9	81-1:95-9	81-1:95-9	81-1:95-9

	GR	NL	PT	FI	UK	US
Bread	81-1:95-12	81-1:95-12	81-1:95-12	85-1:95-5	81-1:95-12	81-1:95-12
Meat	81-1:95-12	81-1:95-12	81-1:95-12	85-1:95-5	81-1:95-12	81-1:95-12
Dairy	81-1:95-12	81-1:95-12	81-1:95-12	85-1:95-5	81-1:95-12	81-1:95-12
Fruits	81-1:95-12	81-1:95-12	81-1:95-12	85-1:95-5	81-1:95-12	81-1:95-12
Tobacco	81-1:95-12	81-1:95-12	81-1:95-12	85-1:95-5	81-1:95-12	81-1:95-12
Alcohol	81-1:95-12	81-1:95-12	81-1:95-11	85-1:95-5	81-1:95-12	81-1:95-12
Clothing	81-1:95-12	81-1:95-12	81-1:95-12	85-1:95-5	81-1:95-12	81-1:95-12
Footwear	81-1:95-12	81-1:95-12	81-1:95-12	85-1:95-5	81-1:95-12	81-1:95-12
Rents	81-1:95-12	81-1:95-12	na	85-1:95-5	81-1:95-12	81-1:95-12
Fuel	81-1:95-12	81-1:95-12	81-1:95-12	na	81-1:95-12	81-1:95-12
Furnit.	81-1:94-10	81-1:94-10	81-1:94-10	na	81-1:94-10	81-1:95-12
Dom. Appl.	81-1:94-10	81-1:94-10	81-1:94-10	na	81-1:94-10	81-1:94-10
Vehicles	81-1:95-12	81-1:95-12	81-1:95-12	85-1:95-5	81-1:95-12	81-1:95-12
Pub. Transp	81-1:95-12	81-1:95-12	81-1:95-12	85-1:95-5	81-1:95-12	81-1:95-12
Comm.	81-1:95-12	81-1:95-12	81-1:95-12	85-1:95-5	81-1:95-12	81-1:95-12
Sound	81-1:95-12	81-1:95-12	na	85-1:95-5	81-1:95-12	81-1:95-12
Leisure	81-1:95-9	81-1:95-9	81-1:95-9	85-1:95-5	81-1:95-9	81-1:95-9
Books	81-1:95-9	81-1:95-9	81-1:95-9	85-1:95-5	81-1:95-9	81-1:95-9
Hotels	81-1:95-9	81-1:95-9	81-1:95-9	85-1:95-5	81-1:95-9	81-1:95-9

University of Lausanne, HEC-FAME and London Business School
Bank of England
European University Institute and Centre for Economic Policy Research
Princeton University, Centre for Economic Policy Research and National Bureau of Economic Research

REFERENCES

Anderson, T.W. and Hsiao, Cheng, "Formulation and Estimation of Dynamic Models using Panel Data," Journal of Econometrics, XVIII (1982), 47-82.

Andrews, Donald W.K., "Exactly Median-Unbiased Estimation of First Order Autoregressive/Unit Root Models," Econometrica, LXI (1993), 139–165.

Andrews, Donald W.K. and Hong-Yuan Chen, "Approximately Median-Unbiased Estimation of Autoregressive Models," Journal of Business and Economic Statistics, XII (II) (1994), 187-204.

Arellano, Manuel and Stephen Bond, "Some Tests of Specification for Panel Data: Monte-Carlo Evidence and an Application to Employment Data," Review of Economic Studies, LVIII (1991), 277-297.

Bergin, Paul and Reuven Glick, "Endogenous Tradability and Macroeconomic Implications," Federal Reserve Bank of San Francisco Working paper no. 2003-09 (2003).

Bils, Mark and Peter Klenow, "Some Evidence on the Importance of Sticky Prices," NBER Working Paper No. 9069 (2002).

Blanchard, Olivier, "Aggregate and Individual Price Adjustment," Brookings Papers on Economic Activity I (1987), 57-109.

Boyd, Derick and Ron Smith, "Testing for Purchasing Power Parity: Econometric Issues and an Application to Developing Countries," Manchester School, LXVII (III) (1999), 287-303.

Brown, Bryan W. and Whitney K. Newey, "GMM, Efficient Bootstrapping and Improved Inference," Journal of Business and Economics Statistics, XX (2002), 507-17.

Campa, Jose and Linda Goldberg, "Exchange Rate Pass-Through into Import Prices: A Macro or Micro Phenomenon?" NBER Working Paper 8934 (2002).

Cashin, Paul and John McDermott, "An Unbiased Appraisal of Purchasing Power Parity," IMF Staff Paper, L (III) (2003,).

Chari, V.V., Kehoe, Patrick and Ellen McGrattan,. "Can Sticky Price Models Generate Volatile and Persistent Real Exchange Rates?" Review of Economic Studies, LXIX(III) (2002), 533–563.

Chen, Shiu-Sheng and Charles Engel, "Does 'Aggregation Bias' Explain the PPP Puzzle?" Pacific Economic Review forthcoming.

Cheung, Yin-Wong, Menzie Chinn and Eiji Fujii, "Market Structure and the Persistence of Sectoral Real Exchange Rates," International Journal of Finance and Economics VI(II) (2001), 95-114.

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

Coakley Jerry , Robert Flood and Mark Taylor, "Long Run Purchasing Power Parity: the First Tests," *Mimeo* IMF (2002).

Crucini, Mario J., Telmer, Chris J. and Marios Zachariadis, "Understanding European Real Exchange Rates," *Mimeo* Vanderbilt University (2001).

Crucini, Mario J. and Mototsugu Shintani, "Persistence in Law-of-One-Price Deviations:

Evidence from Micro-data," *Mimeo* Vanderbilt University (2002).

Diebold, Francis X. and Atsushi Inoue, "Long Memory and Regime Switching," Journal of Econometrics, CV (2002), 131-59.

Engel, Charles, "Long-Run PPP May Not Hold After All," Journal of International Economics, LI (2000a), 243-73.

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

Frankel, Jeffrey A. and Andrew K. Rose, "A Panel Project on Purchasing Power Parity: Mean Reversion Within and Between Countries," Journal of International Economics XL, (1996), 209-25.

Froot, Kenneth and Kenneth S. Rogoff, "Perspectives on PPP and Long-Run Exchange Rates," Handbook of International Economics, Eds. G. Grossman and K. Rogoff, Vol.III (1995), 1647-1688, North-Holland, New York.

Ghironi, Fabio and Marc Melitz, "International Trade and Heterogeneous Firms," NBER working paper no. 10540 (2004).

Giovannini, Alberto, "Exchange Rates and Traded Goods Prices," Journal of International Economics XXIV(I/II) (1988),.45-68.

Goldberg, Pinelopi K. and Michael M. Knetter, "Goods Prices and Exchange Rates: What Have We Learned?" Journal of Economic Literature XXXV (1997), 1243-72.

Granger, Clive, "Long Memory Relationships and the Aggregation of Dynamic Models," Journal of Econometrics, XIV (1980), 227-238.

Granger, Clive and M.J. Morris, "Time Series Modelling and Interpretation," Journal of the Royal Statistical Society A, CXXXIX (1976), 246-257.

Haskel, Jonathan and Holger Wolf, "Why Does the Law of One Price Fail? A Case Study," Scandinavian Journal of Economics, CIII(2001), 545-558.

Hildreth, Clifford and James P. Houck, "Some Estimators for a Linear Model with Random Coefficients," Journal of the American Statistical Association LXIII (1968), 584-95.

Hurwicz, Leonid, "Least Squares Bias in Time Series," in T. Koopmans (ed.): Statistical Inference on Dynamic Economic Models, (1950) New York: Wiley.

Im, Kyung So, M. Hashem Pesaran and Yongcheol Shin, "Testing for Unit Roots in Heterogeneous Panels," DAE Working Paper No. 9526 (1997), University of Cambridge.

Imbs, Jean, Haroon Mumtaz, Morten O.Ravn and Hélène Rey, "Non-Linearities and Real Exchange Rate Dynamics," Journal of the European Economic Association I (2003), 639-49.

Imbs, Jean, Haroon Mumtaz, Morten O.Ravn and Hélène Rey, "Aggregation Bias Does Explain the PPP Puzzle," (2004) available at *http://www.princeton.edu/~hrey/.*

Isard, Peter, "How Far Can We Push the Law of One Price?" American Economic Review LXVI(V) (1977), 942-48.

Kilian, Lutz, "Small-Sample Confidence Intervals for Impulse Response Functions," Review of Economics and Statistics LXXX (II) (1998), 218-30.

Kilian, Lutz and Mark Taylor, "Why is it so Difficult to Beat the Random Walk Forecast of Exchange Rates?", Journal of International Economics, LX(I), (2003), 85-107.

Kilian, Lutz and Zha, Tao, "Quantifying the Uncertainty about the Half-Life of Deviations from PPP," Journal of Applied Econometrics, XVII, II, (2002), 107-125.

Knetter, Michael M., "Price Discrimination by U.S. and German Exporters," American Economic Review, LXXIX(I) (1989), 198-210.

Knetter, Michael M., "International Comparisons of Pricing-to-Market Behavior,." American Economic Review, LXXXIII(III) (1993), 473-86.

Levin, Andrew and Chien-Fu Lin, "Unit Root Tests in Heterogeneous Panels," UCSD Working Paper No. 93-56 (1993).

Lewbel, Arthur and Serena Ng, "Non-Stationary Aggregate Demand Systems and Heterogeneous Consumers," (1993), *Mimeo* University of Michigan.

Lothian, James R., "Multi-Country Evidence on the Behavior of Purchasing Power Parity under the Current Float," Journal of International Money and Finance, XVI(I) (1997) 19-35.

Michael, Panos, A. Robert Nobay and David A. Peel, "Transactions Costs and Nonlinear Adjustment in Real Exchange Rates: An Empirical Investigation," Journal of Political Economy, CV(IV) (1997) 862-79.

Murray, Christian J. and David H. Papell, "The Purchasing Power Parity Persistence Paradigm," Journal of International Economics, LVI (2002a),.1-19.

Murray, Christian J. and David H. Papell, "Do Panels Help Solve the Purchasing Power Puzzle?" (2002b), *Mimeo* University of Houston.

Murray, Christian J. and David H. Papell, "The Purchasing Power Parity Puzzle is Worse Than You Think: A Note on Long-Run Real Exchange Rate Behavior," (2002c), *Mimeo* University of Houston.

Ng, Serena and Pierre Perron, "PPP May not Hold Afterall: A Further Investigation," Annals of Economics and Finance, III (2002), 41-64.

Nickell, Stephen, "Biases in Dynamic Models with Fixed Effects," Econometrica, XLIX (1981), 1417-46.

Oh, Keun-Yeob, "Purchasing Power Parity and Unit Root Tests Using Panel Data," Journal of International Money and Finance, XV(III) (1996),. 405-18.

Obstfeld, Maurice and Alan M. Taylor, "Nonlinear Aspects of Goods-Market Arbitrage and Adjustment: Heckscher's Commodity Points Revisited," Journal of the Japanese and International Economies, XI (1997), 441-79.

Orcutt, Guy H., "A Study of the Autoregressive Nature of the Time Series Used for Tinbergen's Model of the Economic System of the United States," Journal of the Royal Statistical Society (B) (1948), 1-45.

Parsley, David and Shang-Jin Wei, "Convergence to the Law of One Price without Trade Barriers or Currency Fluctuations," Quarterly Journal of Economics (November) (1996), 1211-36.

Parsley, David and Shang-Jin Wei, "A Prism into the PPP Puzzles: The Micro-foundations of Big Mac Real Exchange Rates," (2003), *mimeo* Vanderbilt University.

Pesaran, M. Hashem, "Estimation and Inference in Large Heterogeneous Panels with Cross -Section Dependence," *University of Cambridge* (2002), DAE Discussion Paper no. 0305.

Pesaran, M. Hashem, "Lecture Notes," (2003), *University of Cambridge.*

Pesaran, M. Hashem and Ron Smith, "Estimating Long-Run Relationships From Dynamic Heterogeneous Panels," Journal of Econometrics, LXVIII (1995), 79-113.

Pesaran, M. Hashem and Zhongyun Zhao, "Bias Reduction in Estimating Long-run Relationships from Dynamic Heterogeneous Panels," in C. Hsiao, K. Lahiri, L-F Lee and M.H. Pesaran (eds), Analysis of Panels and Limited Dependent Variables: A Volume in Honour of G. S. Maddala, Cambridge University Press, (1999), Cambridge, chapter 12, 297-321.

Phillips, Peter C.B., and Donggyu Sul, "Dynamic Panel Estimation and Homogeneity Testing Under Cross Section Dependence," Econometrics Journal VI (2003), 217-59.

Pudney, Stephen E., "The Estimation and Testing of Some Error Components Models," (1978), *mimeo* London School of Economics.

Ravn, Morten O., "Imperfect Competition and Prices in Dynamic Trade Model with Comparative Advantage," (2001), *mimeo*, London Business School.

Richardson, J. David, "Some Empirical Evidence on Commodity Arbitrage and the Law of One Price," Journal of International Economics, VIII(II) (1978), 341-51.

Robertson, Donald and James Symons, "Some Strange Properties of Panel Data Estimators," Journal of Applied Econometrics, VII(II) (1992), 175-189.

Rogoff, Kenneth S., "The Purchasing Power Parity Puzzle," Journal of Economic Literature, XXXIV (1996), 647-68.

Rossi, Barbara, "Confidence Intervals for Half-life Deviations from Purchasing Power Parity," (2003) *mimeo* Duke University.

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

Swamy, P.A.B.V., "Efficient Inference in a Random Coefficient Regression Model," Econometrica, XXXVIII (1970), 311-323.

Swamy, P.A.B.V., "Statistical Inference in Random Coefficient Regression Models," Springer-Verlag (1971).

Taylor, Alan M., "Potential Pitfalls for the Purchasing-Power Parity Puzzle? Sampling and Specification Biases in Mean-Reversion Tests of the Law of One Price," Econometrica, LXIX (2001), 473–98.

Taylor, Mark P. and David A. Peel, "Nonlinear Adjustment, Long-Run Equilibrium and Exchange Rate Fundamentals," Journal of International Money and Finance, XIX (2000), 33-53.

Taylor, Mark P., David A. Peel and Lucio Sarno, "Nonlinear Mean-Reversion in Real Exchange Rates: Towards a Solution to the Purchasing Power Parity," International Economic Review, XLII(IV) (2001), 1015-42.

Wu, Yangru, "Are Real Exchange Rates Nonstationary: Evidence from a Panel-Data Test," Journal of Money, Credit, and Banking, XXVIII(I) (1996), 54-63.

Yang, Jaiwen, "Exchange Rate Pass-Through into U.S. Manufacturing Industries," Review of Economics and Statistics, LXXIX (1997), 95-104.

TABLE I

Unit Root Tests

Test	Trend	$\ln\left(\frac{(P_i * e_{i,us})}{P_{us}}\right)$	$e_{i,us}$	$\ln\left(\frac{(P_{ij} * e_{i,us})}{P_{ujs}}\right)$
IPS ADF	no	-2.48050	-2.3638	5.41498
		[0.44723]	[0.009]	[0.000]
IPS ADF	yes	-2.85247	-4.1100	0.31558
		[0.00217]	[0.000]	[0.37616]
LL	no	-2.48050	-4.81754	-5.4429
		[0.0065]	[0.000]	[0.000]
LL	yes	-2.48041	-4.8138	-5.4339
		[0.00656]	[0.000]	[0.000]
LL[1]	no	-40.29696	-6.09430	12.9538
		[0.0000]	[0.000]	[0.000]

P-values are in parentheses. All regressions include an intercept. IPS denotes Im, Pesaran and Shin and LL stands for Levin and Lin. LL[1] is Levin and Lin test that includes individual effects. The lag length for the IPS tests is set to 12.

TABLE II

Persistence Estimates using Aggregate Data

$q_{c,t} = \gamma_c + \sum_{p=1}^{P} \rho_p \; q_{c,t-p} + \varepsilon_t$

Model	P	$\sum_{p=1}^{P} \rho_p$	Half-Life	LAR	CIR
Fixed Effects	18	0.98	46 $(31, 57)$	0.97 $(0.962, 0.981)$	64.38
Anderson-Hsiao	11	0.99	72 $(33, \infty)$	0.96 $(0.941, 1.05)$	109.68
Arrelano Bond	18	0.99	54 (46.75)	0.98 $(0.975, 0.989)$	75.57

${}^{a}H0 : \rho_c = \rho$	-0.4046 (1.0000)	${}^{c}H0 : E(\gamma_c, X) = 0$	25.856 (0.0021)
${}^{b}H0 : \rho_c = \rho$	70.96 (0.9999)	${}^{d}H0 : \gamma_c = 0$	9.8714 (0.000)

The estimates are based on real exchange rates from 11 countries over the period 1981:01-1995:12. The choice of P is based on general to specific lag selection procedure with a maximum lag of 20 for all models, except AH where it was restricted to 12. At each choice of P, the impulse response was examined and the specification was only selected if the IRF was continuous around 0.5. For the GMM estimator two lags of the levels of relative prices were used as instruments.The confidence intervals in the parentheses were estimated using non-parametric bootstrap with 500 replications. Note that the bootstrap for the Arellano and Bond estimator was carried out using the methods described in Brown and Newey [2001]. "LAR" denotes the largest autoregressive root. "CIR" denotes the cumulated impulse response. "a" is the Hausman test for homogeneity, while "b" denotes the Swami test for this hypothesis. "c" and "d" are the Hausman test for random effects and an F-test for fixed effects, respectively.

TABLE III

Persistence Estimates using Disaggregated Data

$q_{ict} = \gamma_c + \sum_{k=1}^{K} \rho_{ik}\ q_{ict-k} + e_{ict}$

Model	P	$\sum_{k=1}^{K} \rho_{ik}$	Half-Life	LAR	CIR
Fixed Effects	12	0.98	36 $(21, 47)$	0.97 $(0.961, 0.981)$	46.71
Fixed Effects (SURE)	12	0.98	34 $(27, 43)$	0.97 $(0.958, 0.978)$	44.30
Fixed Effects (CCE)	12	0.99	58 $(10, 91)$	0.99 $(0.980, 0.995)$	104.20
Mean Group	19	0.97	26 $(14, 28)$	0.95 $(0.903, 0.973)$	33.15
Mean Group (SURE)	20	0.96	22 $(17, 27)$	0.96 $(0.945, 0.968)$	29.48
Mean Group (CCE)	12	0.95	11 $(7, 12)$	0.95 $(0.924, 0.963)$	20.51

$^a H0 : \rho_i = \rho$	98.15 (0.0000)	$^d H0 : E(\gamma_c, X) = 0$	14765 (0.000)
$^b H0 : \rho_i = \rho$	4353.4 (0.0007)	$^e H0 : \gamma_c = 0$	2.1168 (0.000)
$^c H0 : E(\eta_i, X) = 0$	485.02 (0.0022)	$^f LM$	2194698 (0.000)

The estimates are based on relative prices on a maximum of 19 goods from 11 countries over the period 1981:01-1995:12. The choice of P is based on general to specific lag selection procedure with a maximum lag of 20 for all models, except AH where it was restricted to 12. At each choice of P, the impulse response was examined and the specification was only selected if the IRF was continuous around 0.5. The confidence intervals in the parenthesis were estimated using non-parametric bootstrap with 500 replications. "LAR" denotes the largest autoregressive root. "CIR" denotes the cumulated impulse response. "a" is the Hausman test for homogeneity (allowing for correlated residuals), while "b" denotes the Swami test for this hypothesis. "c" is the Pudney [1978] test for the null of no correlation between the random coefficients and the error term. "d" and "e" are the Hausman test for random effects and an F-test for fixed effects, respectively, while "f" is a Breusch-Pagan test for the diagonality of the covariance matrix.

TABLE IV

Persistence Estimates using Disaggregated Data (Bias Corrected)

$q_{ict} = \gamma_c + \sum_{k=1}^{K} \rho_{ik}\ q_{ict-k} + e_{ict}$			
Model	P	Half-Life (Indirect)	Half Life (Direct)
Mean Group	19	41 $(17, 64)$	31 $(17, 57)$
Mean Group (SURE)	5	43 $(18, 105)$	27 $(16, 65)$
Mean Group (CCE)	5	20 $(11, 28)$	18 $(11, 28)$

The Bias Correction is carried out via the Kilian [1998] bootstrap method. "Indirect" refers to a method where ρ is corrected and the half life is estimated on the basis of ρ^*. In the "direct" case, the half-life is corrected directly. In each case, the bootstrap uses 500 replications. For the Mean group model N=204, and T=1981:01 to 1995:12. For the other two models the cross section in Chen and Engel [2004] is used to ensure non-singularity of covariance matrices. In addition, the time series is restricted to 1981:06 to 1994:09 in order to produce a balanced panel. This helps to decrease computation time and has little impact on the underlying (uncorrected) estimates. The confidence intervals are calculated via a double bootstrap procedure. That is, at each replication bootstrap samples are drawn using the mean estimates from the models in the table and the generated data is used to estimate the models via Kilian's bootstrap using 100 replications (50 for Mean Group (SURE)). This is repeated 100 times (50 for Mean Group (SURE)) and the 95% confidence intervals are calculated.

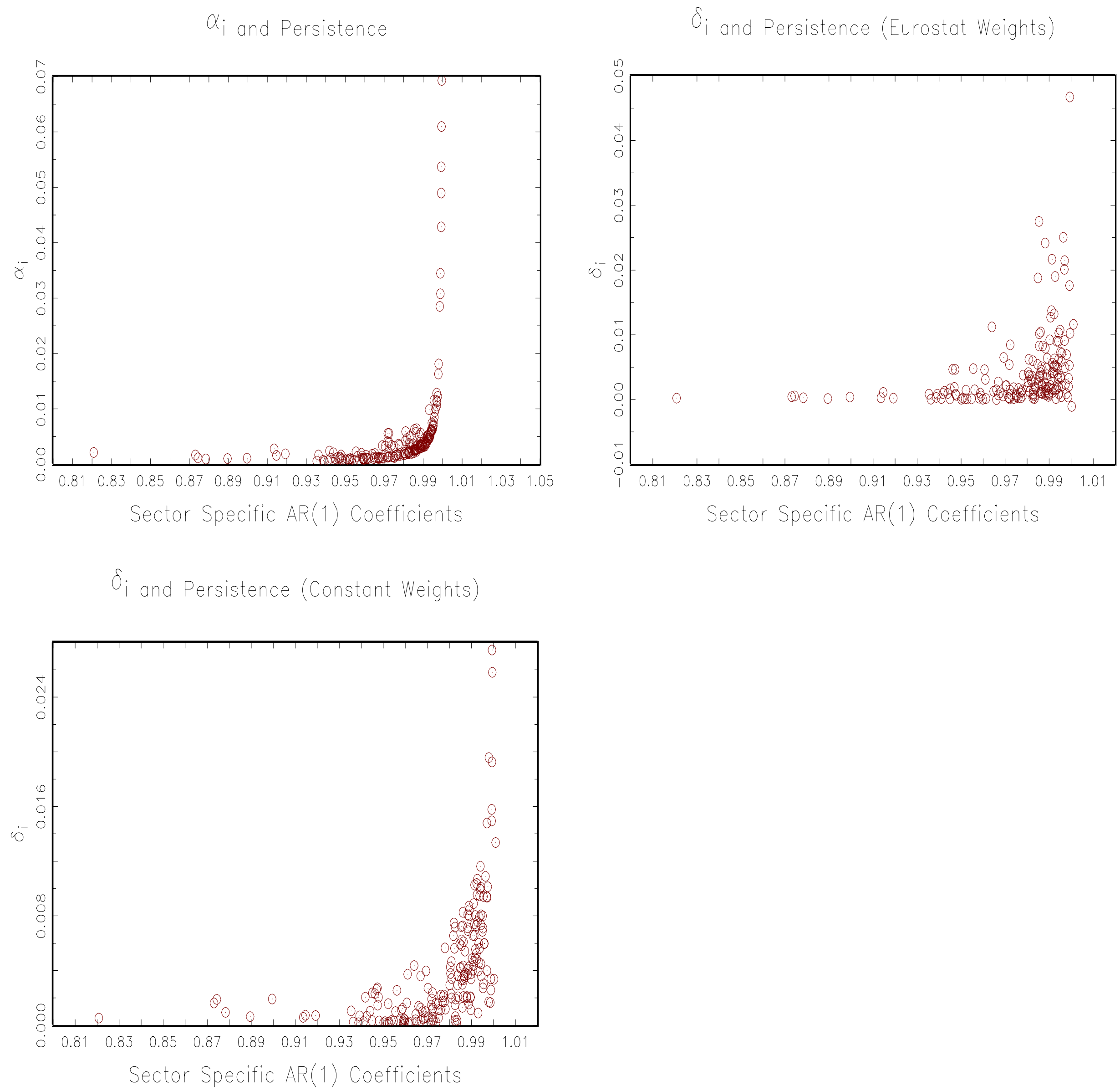

FIGURE I

Check on Conditions for a Positive Aggregation Bias

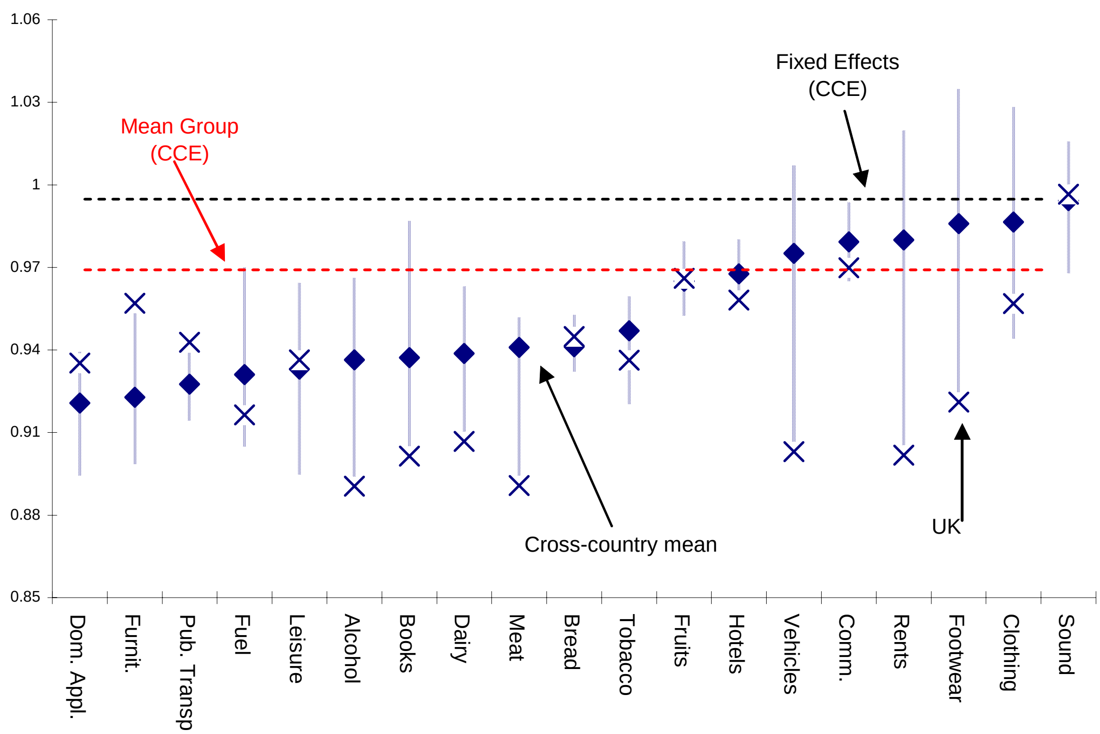

FIGURE II

LAR Estimates

The Figure reports the cross-country mean and inter-quartile range of the MG-CCE estimates for the largest autoregressive roots at the goods level. Actual estimates for the United Kingdom illustrate an example of within-country variation.

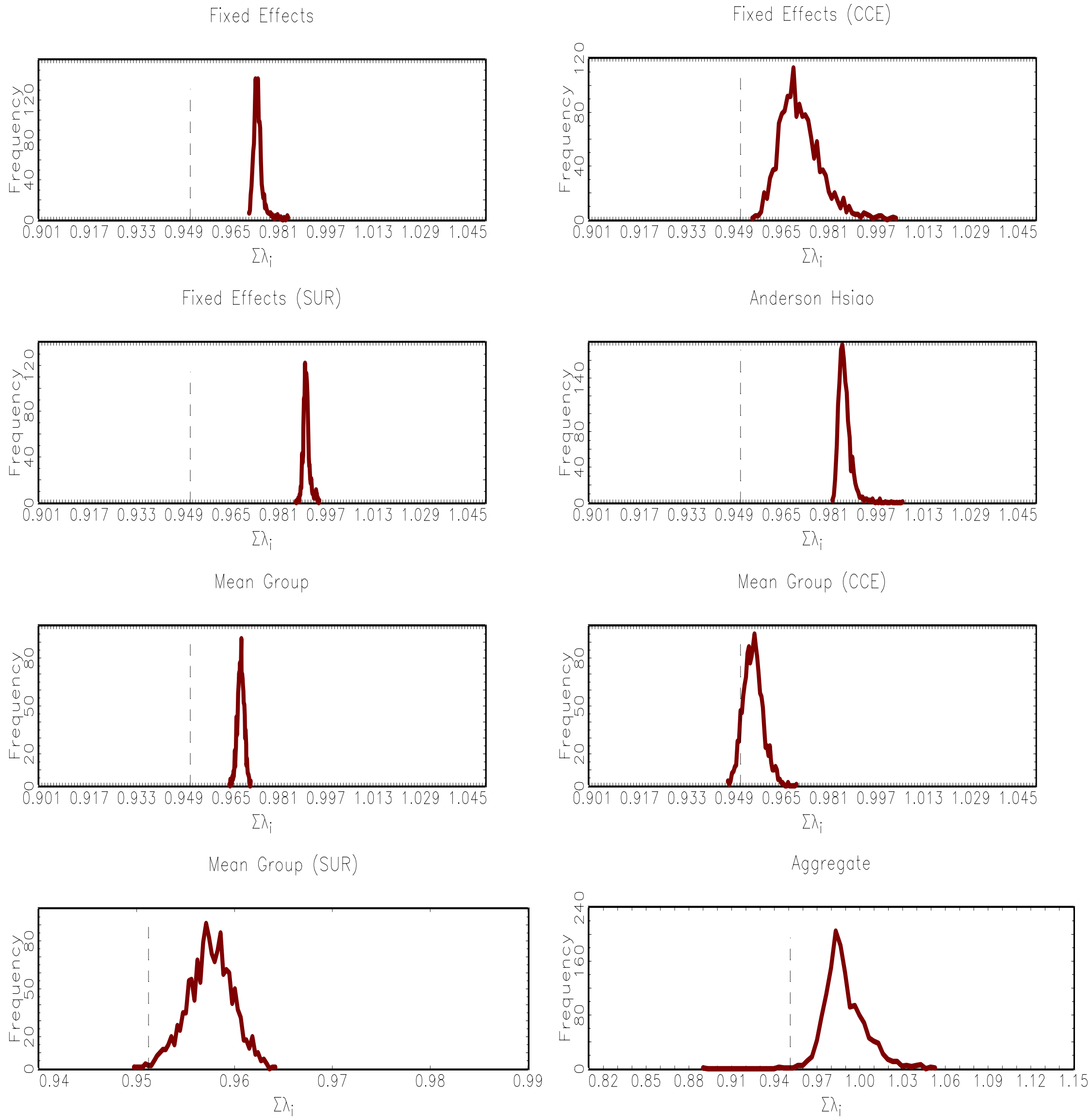

FIGURE III

Monte Carlo Experiment on Estimators

The data was generated using the coefficient estimates and residuals from the Mean group model that allows for the CCE correction. At each replication, iid samples were drawn from each of the N=204 residuals and these (along with the mean $\frac{1}{N}\sum_{i=1}^{N} q_{i,t}$) were used to generate N AR(12) series, using actual observations as starting values and with the correlation structure implied by our coefficient estimates. The generated data was then used to estimate the various models shown above. Note that "aggregate" denotes estimation on time series data obtained by averaging over the cross sections. For the SUR models, the first 60 cross sections were dropped in order to make estimation feasible. The figure plots the histograms of the sum of coefficients obtained from 1500 replications and compares these with the true estimate shown as the dotted black line.

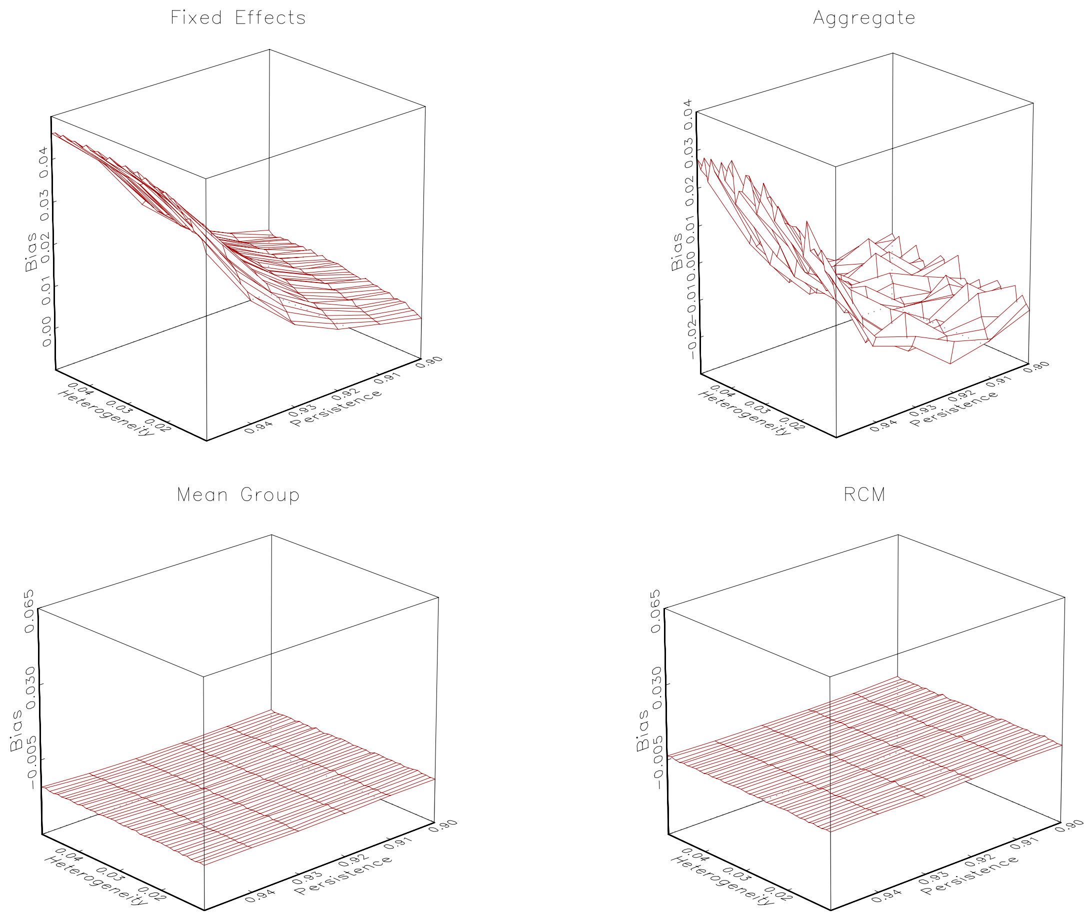

FIGURE IV

The Aggregation Bias

The data was generated form the following AR(1) panel data model (with N=204, T=200): $y_{i,t} = \alpha_i + \lambda_i y_{i,t-1} + v_{i,t}$, where $\alpha_i, v_{i,t} \sim N(0,1)$. The heterogeneous AR coefficients are drawn from the following scheme: $\lambda_i = \lambda + \eta_i$, where $\lambda = \{0.9, 0.91, 0.92, 0.93, 0.94, 0.95\}$ and η_i is sampled from a truncated $N(0, \sigma^2)$ with bounds ± 0.05. We consider 40 values for σ^2 varying from 0.02 to 0.4 with increments of 0.001. This controls the underlying heterogeneity in λ_i and is shown on the "heterogeneity" axis in the figures. We use our data to generate y_0 and then discard the first 50 observations in each cross section to reduce influence of starting values. The figures plot the mean bias in each estimator derived from 200 replications for each combination of λ and σ^2.