

Pairwise Tests of Purchasing Power Parity Using Aggregate and Disaggregate Price Measures

*M. Hashem Pesaran, Ron P. Smith,
Takashi Yamagata and Liudmyla Hvozdyk*

April 2006

CWPE 0634

Pairwise Tests of Purchasing Power Parity Using Aggregate and Disaggregate Price Measures*

M. Hashem Pesaran
Cambridge University

Ron P. Smith
Birkbeck College, London

Takashi Yamagata
Cambridge University

Liudmyla Hvozdyk
Cambridge University

March 2006

Abstract

In this paper we adopt a new approach to testing for purchasing power parity, PPP, that is robust to base country effects, cross-section dependence, and aggregation. We test for PPP applying a pairwise approach to the disaggregated data set recently analysed by Imbs, Mumtaz, Ravan and Rey (2005, QJE). We consider a variety of tests applied to all 66 possible pairs of real exchange rate among the 12 countries and estimate the proportion of the pairs that are stationary, for the aggregates and each of the 19 commodity groups. To deal with small sample problems, we use a factor augmented sieve bootstrap approach and present bootstrap pairwise estimates of the proportions that are stationary. The bootstrapped rejection frequencies at 26%-49% based on unit root tests suggest some evidence in favour of the PPP in the case of the disaggregate data as compared to 6%-14% based on aggregate price series.

JEL Categories: C23, F31, F41

Keywords: Purchasing Power Parity, Panel Data, Pairwise Approach, Cross Section Dependence.

*We are grateful to Yunus Aksoy for useful discussion on this paper. Pesaran and Yamagata gratefully acknowledge the financial support from the ESRC (RES-000-23-0135).

1 Introduction

In this paper we adopt a new approach to examining the purchasing power parity, PPP, hypothesis that is robust to base country effects, cross-section dependence and aggregation. Tests for PPP usually involve examining the time series properties of relative prices in a common currency to determine whether there is adjustment to equilibrium and, if so, to estimate the adjustment speed. Given data on $N + 1$ countries, $i, j = 0, 1, 2, \dots, N$, the standard procedure is to construct the N relative prices against a base country, 0, e.g. the US. The evidence is that the test results can be quite sensitive to the choice of base country. It could be that the real exchange rate between a pair of countries was $I(0)$, but both their rates against the US were $I(1)$. The fact that PPP held between this pair would be lost by just focussing on the relative dollar prices. A closely related issue is that there are usually a high degree of dependence across the different relative prices, which may reflect the base country effect or other common factors, that are difficult to eliminate. Neglect of the cross-section dependence has adverse effects on the properties of estimators and tests and can lead to misleading conclusions. There are also issues of aggregation. Even if individual relative prices adjust quickly, aggregate price indices may not adjust so quickly because the patterns of expenditures in the two countries are very different. In addition heterogeneity in the speeds of adjustment of the individual goods may bias the aggregate or panel estimate of the speed of adjustment towards zero. Imbs, Mumtaz, Ravan and Rey (2005, IMRR), document this heterogeneity bias and for comparability we use their data, which covers 19 goods in 12 countries. We are grateful to them for making their data set available to us.

In this paper we consider a variety of ADF type tests, which have the null hypothesis of no adjustment, and the stationary tests due to Kwiatkowski, Phillips, Schmidt and Shin (1992, KPSS), which have the null of adjustment. We use both aggregate real exchange rates and 19 disaggregate relative prices for particular commodities. Following Pesaran (2005a), who applies this pairwise approach to output and growth convergence, we consider tests using all possible $N(N + 1)/2$ pairs of real exchange rates, or relative prices, between the $N + 1$ countries and estimate the proportion of the pairs that are stationary, for the aggregate and each commodity group separately. This approach is clearly invariant to base country effects and the proportion of country pairs that are non-stationary is consistently estimated despite the cross section dependence of the pairwise tests. To deal with sampling variation in the estimation of the proportions and possible residual cross section dependence, we develop a factor augmented sieve bootstrap procedure, estimating factors by cross-section averages, and present bootstrap pairwise estimates of the proportions that are stationary¹. The main contribution of this paper is empirical; using an existing dataset and an approach recently proposed in the literature, we assess the evidence for PPP in a rather different way. We just present the central results in the paper. A supplement containing a variety of other results is available on request.

There are a range of other issues in the PPP debate that we do not address. Taylor and Taylor (2004) provide a recent survey. These issues include: the relative role of prices and nominal exchange rates to adjustment (e.g. Cheung et al. 2004); non-linear adjustment (e.g. Michael et al. 1997); behavioural models (e.g. De Grauwe and Grimaldi 2006); and the size of the trade costs that may impede price adjustments (e.g. Anderson and van Wincoop 2004, Novy 2005). Most tests of PPP use either nominal exchange rates and price differentials or real exchange rates, as will be done here. Embedding the exchange rate equations in larger models as done by Dees et al. (2006) and Garratt et al. (2006) seems to be more supportive of the PPP hypothesis.

¹Pesaran (2005a) provides a justification for the use of cross-section averages as measures of unobserved factors to deal with cross-section dependence.

Section 2 describes the pairwise approach. Section 3 presents the results. Section 4 bootstraps the proportions. Section 5 contains some concluding comments.

2 The Pairwise Approach

The pairwise approach adopted in this paper addresses different questions from those posed by IMRR and it may be useful to contrast them. IMRR confine their attention to prices relative to the US and, although they note that results relative to the UK were similar, their results may not be robust to base country effects. They emphasise estimation of speed of adjustment, or half life, conditional on adjustment; rather than testing whether there is adjustment. Unlike standard cases, the time series properties of relative prices mean that testing the null hypothesis of no adjustment is not symmetrical with estimating the speed of adjustment and its confidence interval, conditional on adjustment taking place. This is because the distributions under the null, no adjustment, and the alternative, adjustment, are different. Thus the null hypothesis, no adjustment, may not be rejected, yet under the alternative, the confidence interval around the speed of adjustment might not cover zero. They do various panel unit root tests with and without trend concluding that the evidence for stationarity is stronger for the disaggregate data than the aggregate data. However, their tests may reject the null that all the series have a unit root if a proportion of the cases are stationary, even if the proportion that are stationary is not large. Whereas the panel unit root tests are not informative about the proportion that are stationary, the pairwise approach may be, because it estimates the proportion directly.

To deal with cross-section dependence they use the correlated common effect (CCE) estimator of Pesaran (2005b) which can be written:

$$\Delta q_{ict} = \mu_{ic} + \lambda_{ic}(\bar{q}_{t-1} - \theta_{ic}q_{ic,t-1}) + \sum_{k=1}^K \gamma_{ick} \Delta q_{ic,t-k} + \sum_{h=0}^H \delta_{ich} \Delta \bar{q}_{t-h} + e_{ict} \quad (1)$$

where q_{ict} is the relative price for good c in country i relative to US prices in dollars and \bar{q}_t is the cross-section mean of the q_{ict} . For testing PPP, this approach is only valid when it is *known* that the cross-section mean, \bar{q}_t , is stationary. Also the above specification measures the adjustment of q_{ict} to the mean real exchange rate, \bar{q}_t , not to a constant, which is the usual interpretation of PPP. It could be that \bar{q}_t is $I(1)$, $\theta_{ic} \neq 0$ and $\lambda_{ic} > 0$, so that q_{ict} and \bar{q}_t are both $I(1)$ and cointegrated. However, an $I(1)$ real exchange rate is not consistent with the usual interpretation of PPP.

The IMRR data covers monthly price indices for 19 goods, P_{ict} for $c = 1, 2, \dots, C$, e.g. bread, meat, and for aggregate prices P_{it} , ... for the 12, $i, j = 0, 1, 2, \dots, 11$ countries listed in Table 1, with their abbreviations (the US, country 0, and 11 European countries) and exchange rates for country i against the US dollar E_{it} (domestic currency unit per dollar²). Getting consistent cross-country data on disaggregated prices is quite difficult, this is a well documented data set, which has been carefully cleaned and widely used. The samples are mainly from 1981m1 to 1995m12, 180 observations, but a relatively short span. Some series are shorter, in particular the Finland series start from 1985m1 and two series for particular goods in Portugal and three in Finland are missing completely. Full details of the data coverage are given in IMRR Appendix 3. Because the panel is unbalanced, when required we use T_{ijc} to denote the sample size available and N_c the number of countries available for a particular good, c .

Given data on E_{it} , the nominal bilateral exchange rate between country i and the US; P_{it} the aggregate price index for country i at month t ; and P_t the log of the US price index; then

²IMRR use dollars per domestic currency unit, but we converted it to follow the usual convention.

the log real exchange rate for country i relative to the US is:

$$q_{it} = \ln\left(\frac{E_{it}P_t}{P_{it}}\right) = e_{it} + p_t - p_{it}, \quad (2)$$

using lower case letters to denote logarithms. Since $E_{ijt} = E_{it}/E_{jt}$, the real exchange rate between any other pair of countries $i, j \neq 0$ can be calculated as

$$q_{it} - q_{jt} = \ln\left(\frac{E_{ijt}P_{jt}}{P_{it}}\right) = q_{ijt}. \quad (3)$$

We will refer to these q_{ijt} as pairwise real exchange rates, between i and j , where we will include the US in the pairs, namely consider all $i, j = 0, 1, \dots, N$, for $i \neq j$. Similarly, given data on P_{ict} and P_{jct} , the price indices for good c in country i and j at date t , we can form the log relative prices for the good $c = 1, 2, \dots, C$ for country i relative to country j as:

$$q_{ict} = \ln\left(\frac{E_{it}P_{ct}}{P_{ict}}\right), \quad (4)$$

$$q_{ijct} = q_{ict} - q_{jct}. \quad (5)$$

We consider tests using all possible $N(N+1)/2$ distinct pairs of real exchange rate q_{ijt} or relative prices q_{ijct} amongst the $N+1$ countries and estimate the proportion of the pairs that are stationary, for the aggregate and each good, using a variety of tests. As argued in Pesaran (2005a), the average rejection rate is likely to be more robust to cross-section dependence and the possibility of an $I(1)$ unobserved factor inducing that dependence than the alternative methods available.

Consider the following factor model for the dollar real exchange rates:

$$q_{it} = \alpha_i + \gamma_i' \mathbf{f}_t + \varepsilon_{it}. \quad (6)$$

There is an $I(0)$ idiosyncratic component, ε_{it} and the common factors, \mathbf{f}_t , which induce cross-sectional dependence and may be $I(0)$ or $I(1)$. Then

$$q_{ijt} = (\alpha_i - \alpha_j) + (\gamma_i - \gamma_j)' \mathbf{f}_t + \varepsilon_{it} - \varepsilon_{jt}, \quad (7)$$

will be $I(0)$ if either \mathbf{f}_t is $I(0)$ or if \mathbf{f}_t is $I(1)$ and non-cointegrated but $\gamma_i = \gamma_j$.

Consider an ADF test, where the null is a unit root, non-stationarity, H_N , with lag order p , applied to q_{ijt} , $t = 1, 2, \dots, T$. Let $Z_{ijt} = 1$ if $ADF_{ijt}(p) < K_{T,p,\alpha}$ the critical value for an ADF test of size α , such that $\lim_{T \rightarrow \infty} \Pr(ADF_{ijt}(p) < K_{T,p,\alpha} \mid H_N) = \alpha$. The fraction of the $N(N+1)/2$ pairs for which the unit root null is rejected is

$$\bar{Z}_{NT}^A = \frac{2}{N(N+1)} \sum_{i=0}^{N-1} \sum_{j=i+1}^N Z_{ijt}. \quad (8)$$

Even if the idiosyncratic components ε_{it} are independent across countries Z_{ijt} and Z_{ikt} will not be independently distributed. However, Pesaran (2005a) shows that if the idiosyncratic components are independent, under H_N , \bar{Z}_{NT}^A is a consistent estimator of α for large N and T . Under the stationary alternative, H_S , \bar{Z}_{NT}^A converges to unity for large N and T . Similar results apply to the KPSS test with the role of the null and alternative reversed. In that case \bar{Z}_{NT}^K is the proportion of cases where the stationary hypothesis, H_S , is rejected.

If H_S holds then we would expect \overline{Z}_{NT}^K to be close to the chosen size of the test and \overline{Z}_{NT}^A to be large. If H_N holds then we would expect \overline{Z}_{NT}^K to be large and \overline{Z}_{NT}^A to be close to the size of the test. When T is finite the proportions \overline{Z}_{NT}^K converges to α_T , the empirical rejection frequency of the underlying test for the appropriate null. The average rejection frequency converges to α_T as $N \rightarrow \infty$. Increasing the panel dimension reduces the sampling variation of the estimated proportions. A similar reduction in sampling variation can be achieved by increasing the number of replications in a bootstrap experiment, a point we return to below. In principle, it would be possible to develop a formal statistical test of whether the estimated proportion of rejections is significantly greater than the size of the test. In practice, the magnitude of the proportion of rejections is of more economic interest.

As a final preliminary it is worth considering the relation between the CCE estimator, the pairwise approach, and tests of PPP using effective real exchange rates. The effective real exchange rate is defined

$$\bar{q}_{it} = \sum_{j=0}^N w_{ij} q_{ijt} \quad (9)$$

where w_{ij} is the share of country j in the trade (exports plus imports) of country i . This is closely related to the \bar{q}_t used as a regressor in the CCE estimator by IMRR, which uses equal weights rather than trade weights. It is also a weighted average of the pairwise real exchange rates for country i . The effective real rate of country i will be $I(0)$ if all the pairwise rates of country i are $I(0)$ (given non-zero trade weights that add up to unity). The effective real rate will be $I(1)$ if some of the pairwise rates are $I(1)$.

To investigate their properties, effective real rates were constructed for the twelve countries in the sample using data on imports and exports from the IMF Direction of Trade Statistics. The weights w_{ij} are the share for country j in the total trade of country i with all 11 countries in the sample; using average trade flows over the period 1987-1989, about the middle of the estimation period. Details of their construction and full results are given in section 1 of a supplement that is available on request. Table 2 summarises the results. It gives ADF statistics, both for the case with just a constant and the case with constant and linear trend, where lag length is selected by the AIC for a maximum of lag of 12. As one would expect from the literature, there is little evidence against the unit root hypothesis. Belgium rejects at the 5% level in the intercept and trend case, the UK at the 10% level in the intercept case. However, as noted above, this only indicates that not all the pairwise rates are $I(0)$. Therefore, we investigate the pairwise rates directly.

3 Results of the Pairwise Tests

Conducting the tests requires a number of choices. We make these choices in a way that is likely to be favourable to the PPP hypothesis, which is the hypothesis of interest. We assume that the equilibrium real exchange rate is given by an intercept and trend. Results with just an intercept are given in section 2 of the supplement. They showed fewer rejections of the unit root null as one might expect. We use the standard ADF, the ADF-GLS of Elliot et al. (1996), and the ADF-WS of Park and Fuller (1995). All three have the null of a unit root. The ADF-GLS and ADF-WS are designed to have higher power than the standard ADF. We also use the KPSS (1992) test, which has the null of stationarity. Whichever test we use, we will present the results as the proportion of cases that reject the null of that test. Since the data are monthly, we examined both the seasonally adjusted and unadjusted price series. Seasonal adjustment did not make a lot of difference to the results. Details on the method of seasonal adjustment and the results

on seasonally adjusted data are given in section 3 of the supplement. In ADF type tests, we allow for a lag length, p , up to 12 and use both AIC and SBC to select lag length. For KPSS tests we use three window sizes: $0.75T^{1/3}$, which is standard in econometric applications; the much larger window size of $2T^{1/2}$, often used in the time series literature; and the average of the two. The KPSS test relies on estimating the spectral density at zero frequency to correct for serially correlated errors, which is subject to a considerable degree of uncertainty. To avoid size distortions one needs a large window size, but this is likely to reduce power. Considering a number of window sizes allows us to investigate this trade off. We consider the full sample of 12 countries, with a maximum of 66 pairwise comparisons across a maximum of 19 goods (data on prices of some goods for some countries do not exist). We use a 10% size. The appropriate critical values for each of the ADF type tests were calculated by Monte Carlo using 10,000 replications for each p and T . The critical values are given in section 4 of the supplement.

The fraction of pairs (out of a maximum of 66) that reject the null hypothesis for the aggregate CPI and the 19 subcategories of goods prices are given in Table 3. The results are very sensitive to treatment of dynamics, much more so than output, considered in Pesaran (2005a). This is not surprising, given that the PPP data are monthly, whereas the output series are annual. The results are also sensitive to the model selection criterion used to determine the lag length and to the particular unit root test being considered. The rejection rates can differ substantially between the six ADF type procedures (3 tests, 2 methods to select lag length). For instance, rejection rates range from 1.82% to 16.36% for rents and 22.73% to 68.18% for fruits. In 13 of the 20 cases ADF-AIC had the highest fraction of rejections; in one case jointly with one other procedure and in two cases jointly with two other procedures. This is somewhat surprising given that the ADF-AIC is not the procedure one might have expected to have the highest power. The ADF-WS-AIC never has the maximum rejection rate, but the ADF-WS-SBC has the maximum rejection rate in four cases. For the ADF-GLS and the ADF-WS the SBC variant has higher rejection frequencies than the AIC variant. In the case with just an intercept, reported in section 2 of the supplement, ADF-WS has the highest rejection rates and it seems that including a trend reduces the power of the ADF-WS relative to the other tests. The rejection rate is below the size, 10%, in 28 of the 120 cases and above 20% in 30 of the 120 cases. Overall, although the rejection rates are not high, they are very unlikely to have been generated by a process in which all the relative prices had unit roots.

The pattern noted by IMRR is apparent: the fraction of rejections by all three ADF tests is higher on average for the commodities than it is for the aggregate CPI. Whereas the proportion of rejections using the aggregate CPI varies from 6% to 14%, depending on test; the averages over the 19 categories range from 13% to 21%. This is also true in the case with just an intercept. The strongest case for PPP is fruits. When the lag length is chosen by the SBC all three ADF tests reject the unit root in fruits prices in over half the cases. However, when the lag length is chosen by the AIC, the rejection rate is smaller. This pattern can also be seen in the case with just an intercept: when SBC is used the unit root is rejected in around half the cases, but this proportion is much lower using AIC. Average lag length and degree of serial correlation seem to differ by commodity. Details are given in section 5 of the supplement.

The KPSS test shows a similar sensitivity to dynamic specification. When the window is set to the standard size of $0.75T^{1/3}$, the null of stationarity is rejected over 95% of the time in all cases, except for fruit, tobacco and domestic appliances. For two commodities the stationary null is rejected in 100% of the cases. However, with the larger window size of $2T^{1/2}$, the rejection rate falls sharply as one would expect and stationarity is only rejected a quarter of the pairwise tests, on average. Using a window size which is the average of the two, gives a rejection rate of 55%. In the case with just an intercept, rejection rates are slightly lower with the standard window and slightly higher with the large window. Asymptotic critical values are used for the

KPSS test, so there may be more risk of size distortion.

Our results confirm IMRR's conclusion that the disaggregate data on individual commodities provide more evidence for adjustment to PPP than the aggregate CPI data. In addition the proportion of rejections is substantially higher than the size of the test, providing quite strong evidence against unit roots in all series. The fact that the proportion of rejections is still quite small may reflect the low power of these tests. Although $T = 180$ is quite large, the actual time span covered is relatively short and the dynamics complicated. In small samples the ADF critical values are sensitive to p and T and in practice p is unknown. In addition, there may be sampling variation in the estimated rejection frequencies because N , the number of countries, is small. To try to improve the small sample properties we use a factor augmented sieve bootstrap. This involves estimating a model which captures the factor structure and serial correlation patterns of the data, then bootstrapping the rejection rates.

4 Factor Augmented Sieve Bootstrap Estimates

It is now standard in the literature to interpret the cross-section dependence in terms of a factor model and such a model is the basis of the CCE estimation approach utilized by IMRR. Therefore we estimate the parameters of the underlying factor model directly and use these estimates to bootstrap the pairwise rejection rates. Whereas in some contexts the factors themselves are of interest, here they are nuisance parameters which induce cross-section dependence, but need to be controlled for if we are to get good estimates of the proportion of rejections.

4.1 Factor Augmented Model

The first stage is to estimate time varying common factors, at the aggregate level (one) and at the commodity levels (nineteen), as:

$$\bar{q}_{ct} = \sum_{i=0}^{N_c} q_{ict} / (N_c + 1), \text{ for } c = 1, 2, \dots, C; \bar{q}_t = \sum_{c=1}^C \bar{q}_{ct} / C. \quad (10)$$

Using equal weights corresponds to those used in the CCE estimator, though estimates of factors tend not to be very sensitive to choice of the weights. Then $ADF(s_c)$ regressions are estimated for \bar{q}_{ct} ,

$$\Delta \bar{q}_{ct} = \hat{\mu}_c + \hat{\phi}_c \bar{q}_{c,t-1} + \sum_{\ell=1}^{s_c} \hat{\xi}_{c\ell} \Delta \bar{q}_{c,t-\ell} + \hat{e}_{ct}, \quad (11)$$

where the lag-order s_c is chosen by AIC. The null of non-stationarity could not be rejected for \bar{q}_t or any of the \bar{q}_{ct} . The results of ADF tests are given in Table 4. This suggests that the common factors used in the IMRR CCE estimators may well be $I(1)$. Where $\hat{\phi}_c > 0$, we set $\hat{\phi}_c = 0$ and $\hat{\mu}_c = 0$ to avoid explosive factors. Given the uncertainty about whether there are $I(1)$ common factors, we construct the bootstrap under two assumptions: (a) not imposing unit roots on the factors and using the estimated $\hat{\phi}_c$ or (b) imposing unit roots on the factors on the basis of the pretest results and setting $\hat{\phi}_c = 0$, for all c . This allows us to assess the effect of any downward T_c -bias in $\hat{\phi}_c$.

In the second stage, q_{ict} is regressed on unity, \bar{q}_{ct} and \bar{q}_t for each combination of (i, c) to obtain the estimated model

$$q_{ict} = \hat{\alpha}_{ic} + \hat{\gamma}_{ic1} \bar{q}_{ct} + \hat{\gamma}_{ic2} \bar{q}_t + \hat{e}_{ict}. \quad (12)$$

The coefficients on the factors are often large, elasticities of one or greater and there is some tendency for the two factors to have opposite signs. Estimates are given in section 6 of the supplement. There is also a tendency for the neighbouring European countries to have similar values for the factor coefficients. A number of pairs of European countries have goods where coefficients of both factor loadings differ by 0.1 or less. Details of these pairs are given in section 7 of the supplement.

To allow for the serial correlation in $\hat{\varepsilon}_{ict}$, ADF(s_{ic}) regressions are estimated for $\hat{\varepsilon}_{ict}$, for each (i, c) ,

$$\Delta \hat{\varepsilon}_{ict} = \hat{\eta}_{ic} + \hat{\lambda}_{ic} \hat{\varepsilon}_{ic,t-1} + \sum_{\ell=1}^{s_{ic}} \hat{\psi}_{ic\ell} \Delta \hat{\varepsilon}_{ic,t-\ell} + \hat{v}_{ict}, \quad (13)$$

where the lag-order s_{ic} are chosen by AIC. Again, when $\hat{\lambda}_{ic} > 0$, we restrict $\hat{\lambda}_{ic} \equiv 0$ and $\hat{\eta}_{ic} \equiv 0$, to avoid explosive outcomes.

4.2 The Bootstrap Procedure

In the third stage, data are generated on q_{ict} based on the estimates of equations (11), (12), and (13). The residuals of (11) and (13) are drawn with replacement.

Step 1: (a) The r^{th} replication of commodity specific common factor, $\bar{q}_{ct}^{(r)}$, and their cross section average $\bar{q}_t^{(r)}$, are generated as

$$\bar{q}_{ct}^{(r)} = \hat{\mu}_c + (1 + \hat{\phi}_c) \bar{q}_{c,t-1} + \sum_{\ell=1}^{s_c} \hat{\xi}_{c\ell} \Delta \bar{q}_{c,t-\ell} + e_{ct}^{(r)}, \quad t = 1, 2, \dots, T; c = 1, 2, \dots, C \quad (14)$$

$$\bar{q}_t^{(r)} = C^{-1} \sum_{c=1}^C \bar{q}_{ct}^{(r)}, \quad t = 1, 2, \dots, T,$$

where $r = 1, 2, \dots, R$, estimates and lag-orders s_c is defined in (11), $e_{ct}^{(r)}$ are random draws with replacement from $\{\hat{\varepsilon}_{ct}\}_{t=1}^T$, using the initialisations $(\bar{q}_{c,-s_c}^{(r)}, \bar{q}_{c,-(s_c-1)}^{(r)}, \dots, \bar{q}_{c,0}^{(r)}) = (\bar{q}_{c,1}, \bar{q}_{c,2}, \dots, \bar{q}_{c,s_c+1})$ are used.

(b) Step 1(a) uses the estimated $\hat{\phi}_c$ but pretesting indicated that one could not reject the hypothesis $\hat{\phi}_c = 0$, therefore the second set of $\bar{q}_{ct}^{(r)}$, are generated imposing unit roots on commodity specific factors using

$$\bar{q}_{ct}^{(r)} = \bar{q}_{c,t-1} + \sum_{\ell=1}^{s_c} \hat{\xi}_{c\ell} \Delta \bar{q}_{c,t-\ell} + e_{ct}^{(r)}, \quad t = 1, 2, \dots, T; c = 1, 2, \dots, C. \quad (15)$$

the steps below are followed for both measures of $\bar{q}_{ct}^{(r)}$

Step 2: The r^{th} replication of q_{ict} is generated as

$$q_{ict}^{(r)} = \hat{\alpha}_{ic} + \hat{\gamma}_{ic1} \bar{q}_{ct}^{(r)} + \hat{\gamma}_{ic2} \bar{q}_t^{(r)} + \varepsilon_{ict}^{(r)}, \quad i = 1, 2, \dots, N; t = 1, 2, \dots, T; r = 1, 2, \dots, R, \quad (16)$$

where estimates are obtained from (12), and $\varepsilon_{ict}^{(r)}$ are generated using (13) such that

$$\varepsilon_{ict}^{(r)} = \hat{\eta}_{ic} + (1 + \hat{\lambda}_{ic}) \varepsilon_{ic,t-1} + \sum_{\ell=1}^{s_{ic}} \hat{\psi}_{ic\ell} \Delta \varepsilon_{ic,t-\ell} + v_{ict}^{(r)}, \quad (17)$$

where $\hat{\eta}_{ic}$, $\hat{\lambda}_{ic}$, $\hat{\psi}_{icl}$, s_{ic} are estimated in (13), $v_{ict}^{(r)}$ are random draws with replacement from $\{\hat{v}_{ict}\}_{t=1}^T$, initialisations $(\varepsilon_{ic,t-s_{ic}}^{(r)}, \varepsilon_{ic,t-(s_{ic}+1)}^{(r)}, \dots, \varepsilon_{ic0}^{(r)}) = (\hat{\varepsilon}_{ic,1}, \hat{\varepsilon}_{ic,2}, \dots, \hat{\varepsilon}_{ic,s_{ic}+1})$ are used.

Step 3: For each commodity category, we calculate the fraction of the pairs $q_{ijct}^{(r)} = q_{ict}^{(r)} - q_{jct}^{(r)}$ for which the null hypothesis is rejected by the test, $F_c^{(r)}$.

Step 4: Repeat Steps 1 to 3, R times, then obtain the average of the fractions over replications

$$R^{-1} \sum_{r=1}^R F_c^{(r)}.$$

The ADF-GLS and ADF-WS versions are done similarly. We set $T = 180$ (balanced), $N_c = 12$, $C = 19$ as in the full sample.

4.3 Bootstrapped Results

The bootstrapped rejection rates for the case where unit roots were not imposed on the commodity factors are given in Table 5; those where the unit root was imposed on the factors are given in Table 6. As noted above, imposing a unit root on the common commodity factors does not necessarily impose a unit root on the cross-rates. Since

$$\begin{aligned} q_{ijct}^{(r)} &= q_{ict}^{(r)} - q_{jct}^{(r)} \\ &= (\hat{\alpha}_{ic} - \hat{\alpha}_{jc}) + (\hat{\gamma}_{ic1} - \hat{\gamma}_{jc1})\bar{q}_{ct}^{(r)} + (\hat{\gamma}_{ic2} - \hat{\gamma}_{jc2})\bar{q}_t^{(r)} + \varepsilon_{ict}^{(r)} - \varepsilon_{jct}^{(r)}, \end{aligned}$$

then if the factor loadings are similar, $(\hat{\gamma}_{ic1} - \hat{\gamma}_{jc1}) \approx (\hat{\gamma}_{ic2} - \hat{\gamma}_{jc2}) \approx 0$, the I(1) component drops out. As noted above, this seems to be the case for a number of European pairs. In addition, whereas $\varepsilon_{ict}^{(r)}$ and $\varepsilon_{jct}^{(r)}$ follow exact autoregressions by construction, this need not be the case for the difference, $\varepsilon_{ict}^{(r)} - \varepsilon_{jct}^{(r)}$.

As one might expect, the rejection fractions for ADF type tests are higher for the case where unit roots are not imposed on the factors than for the case where unit roots are imposed. However, bootstrapped rejection rates for both cases are higher than the rejection rates for the original data given in Table 3. Not imposing unit roots gives a range for the rejection frequencies of 36% to 49%, imposing unit roots, 26% to 32%. Many features of the original data are maintained, e.g. fruits has a very high rejection rate and there are positive correlations between the actual and bootstrapped rejection rates. The dispersion among the six procedures is much smaller in the bootstrapped data than in the original data: one does not get such large differences in rejection frequencies between tests. This may be because of the reduction in sampling variation. Also unlike the original data, where the standard ADF had the highest average rejection frequency, with the bootstrapped data the ADF-GLS and ADF-WS have higher rejection frequencies than the standard ADF, thus being consistent with the expected power characteristics. With the original data, ADF-AIC had slightly higher rejection frequencies than ADF-SBC, while for the other two tests the SBC variant had slightly higher rejection frequencies. With the bootstrapped data, the AIC variant always has higher rejection frequencies than the SBC variant.

With the KPSS test the position is not so straightforward. Using the standard window, the proportion of rejections of stationarity is much lower in the bootstrap data than the original data, consistent with the ADF results. However, with the large window the proportion of rejections of stationarity is higher with the bootstrapped data than with the original data.

5 Conclusions

In this paper we suggest a new procedure for testing PPP, which avoids a number of difficulties associated with the procedures used by IMRR, whose data we use. We test the null hypothesis of no adjustment towards PPP using a variety of ADF type tests and the null hypothesis of adjustment using KPSS tests with a variety of window sizes. We use both aggregate CPI data and 19 disaggregate relative prices for individual goods and test all possible $N(N + 1)/2$ pairs of real exchange rate pairs across the $N + 1$ countries. This allows us to estimate the proportion of the pairs that reject the null for the particular test. This approach is invariant to base country effects and the proportion that are non-stationary is consistently estimated under the null hypothesis despite the cross-section dependence of the pairwise tests. The rejection rates on ADF type tests are higher than the size of the test, which is the rejection frequency one would expect if all the relative prices had unit roots. Thus the results provide some evidence in favour of PPP. In line with IMRR's findings, the rejection rates are higher on average for the individual commodity prices than for the aggregate price indices. The rejection rates seem sensitive to dynamic specification and because the number of countries is relatively small they are subject to sampling variability. In addition, the rejection rates of the different types of ADF test do not match their known power characteristics. The results from the KPSS test are inconclusive because they are so sensitive to the lag-window used, and there may be size distortions because of the use of asymptotic critical values.

To deal with small sample problems and possible residual cross section dependencies, we estimate the parameters of a factor model from the data and bootstrap the rejection rates. We estimate the global and commodity specific factors that induce the cross-sectional dependence. We cannot reject unit roots in these factors so we conduct the analysis both by not imposing unit roots in the factors and using the estimated coefficients, and by imposing unit roots in the factors. The bootstrapped estimates of the rejection frequencies are more stable, showing less variation between types of test, and the rejection frequencies of the different procedures match their known power characteristics. Also the bootstrapped rejection frequencies of ADF type tests are higher than those obtained from the original data. In the original data the averages over commodities ranged from 13% to 21% depending on unit root test, for the bootstrapped data they range from 36% to 49% when a unit root is not imposed on the factors and 26% to 32% when a unit root is imposed on the factors. These results provide evidence against the null of no PPP and evidence that rejection of this null is stronger with disaggregate data. It also suggests that sampling variability may cause a downward bias in the rejection frequencies. Of course, despite its advantages, this procedure cannot compensate for the fact that the short span of the data used may lower the power of the test. However, long-span data at a disaggregate level is not available and there are issues of potential structural instability. Nevertheless, the pairwise procedure does deal with base-country effects, aggregation, and by using the factor augmented sieve bootstrap procedure it can also potentially deal with the small sample effects, and possible residual cross section dependencies.

References

- Anderson, J. and van Wincoop, E., 2004. Trade costs. *Journal of Economic Literature* 42, 691–751.
- Cheung, Y-W, Lai, K.S., Bergman, M., 2004. Dissecting the PPP puzzle: the unconventional roles of nominal exchange rate and price adjustments. *Journal of International Economics* 64, 135–50.
- Dees, S., di Mauro, F., Pesaran, M.H., Smith, L.V., 2006. Exploring the international linkages of the Euro area: A global VAR analysis. mimeo Cambridge, forthcoming *Journal of Applied Econometrics*.
- De Grauwe, P., Grimaldi, M., 2006. Exchange rate puzzles: A tale of switching attractors. *European Economic Review* 50, 1–33.
- Elliot, G., Rothenberg, T.J., Stock, J.H., 1996. Efficient tests for an autoregressive unit root. *Econometrica* 64, 813–836.
- Garratt, A., Lee, K., Pesaran, M.H., Shin, Y., 2006. Global and National Macroeconometric Modelling: A long-run structural approach. forthcoming Oxford University Press.
- Imbs, J., Mumtaz, H., Ravn, M.O., Rey, H., 2005. PPP strikes back: Aggregation and the real exchange rate. *Quarterly Journal of Economics* 120, 1–43.
- Kwiatkowski, D., Phillips, P.C.B., Schmidt, P., Shin, Y., 1992. Testing the null hypothesis of stationarity against the alternative of a unit root. *Journal of Econometrics* 54, 159–178.
- Michael, P., Nobay, A.R., Peel, D.A., 1997. Transactions costs and non linear adjustment in the real exchange rate. *Journal of Political Economy* 105, 862–79.
- Novy, D., 2005. Is the iceberg melting less quickly? International trade costs after World War II. Mimeo, Cambridge.
- Park, H.J., Fuller, W.A., 1995. Alternative estimators and unit root tests for the autoregressive process. *Journal of Time Series Analysis* 16, 449–459.
- Pesaran, M.H., 2005a. A pair-wise approach to testing for output and growth convergence. forthcoming *Journal of Econometrics*.
- Pesaran, M.H., 2005b. Estimation and inference in large heterogeneous panels with a multifactor error structure, forthcoming *Econometrica*.
- Taylor, A., Taylor, M., 2004. The purchasing power parity debate. *Journal of Economic Perspectives* 18, 135–158.

Table 1: Country Codes

BE	Belgium
DE	Germany
DK	Denmark
ES	Spain
IT	Italy
FR	France
GR	Greece
NL	Netherlands
PT	Portugal
FI	Finland
UK	United Kingdom
US	United States of America

Table 2: ADF Test Results for Real Effective Exchange Rates

	ADF statistics	
	Intercept	Intercept & Trend
BE	-1.602	-5.382**
DE	-0.777	-1.716
DK	-1.325	-2.465
ES	-1.257	-1.602
IT	-1.266	-1.193
FR	-2.270	-2.936
GR	-1.600	-2.125
NL	-1.520	-1.628
PT	-0.549	-2.657
FI	-1.320	-1.501
UK	-2.641*	-2.914
US	-1.245	-2.414

Notes: Reported statistics are based on ADF regressions with maximum AIC values, using a maximum lag of 12. The first column is the case with just an intercept, the second with an intercept and a linear trend. Superscript * signifies the null hypothesis of a unit root is rejected at 10% level. ** signifies the null hypothesis of a unit root is rejected at 5% level. The critical values are simulated with 10,000 replications, based on ADF regressions with the same lag order and sample size as of the underlying ADF regression based on actual data.

Table 3: Fractions of Pairs of q_{ijct} for Which the Null Hypothesis is Rejected at 10% Significance Level - With an Intercept and Linear Trend, for all 12 Countries

	Lag Order Chosen by AIC			Lag Order Chosen by SBC			KPSS		
	ADF	ADF-GLS	ADF-WS	ADF	ADF-GLS	ADF-WS	$0.75T^{1/3}$	$2T^{1/2}$	av
CPI	13.64	6.06	9.09	13.64	13.64	12.12	96.97	31.82	60.61
Subcategories									
Bread	13.64	13.64	12.12	13.64	7.58	10.61	98.48	22.73	53.03
Meat	21.21	18.18	19.70	16.67	19.70	18.18	98.48	28.79	65.15
Dairy	16.67	12.12	13.64	15.15	10.61	13.64	95.45	34.85	62.12
Fruits	28.79	22.73	31.82	59.09	51.52	68.18	90.91	24.24	56.06
Tobacco	40.91	15.15	22.73	42.42	19.70	25.76	93.94	16.67	56.06
Alcohol	19.70	16.67	18.18	15.15	12.12	13.64	96.97	28.79	66.67
Clothing	25.76	18.18	16.67	22.73	15.15	15.15	96.97	18.18	46.97
Footwear	21.21	6.06	7.58	16.67	4.55	4.55	95.45	28.79	50.00
Rents	16.36	7.27	10.91	12.73	1.82	3.64	98.18	34.55	52.73
Fuel	27.27	23.64	27.27	32.73	30.91	34.55	96.36	30.91	61.82
Furnit.	27.27	5.45	14.55	21.82	5.45	10.91	96.36	21.82	41.82
Dom. Appl.	27.27	10.91	16.36	14.55	5.45	9.09	94.55	12.73	36.36
Vehicles	21.21	18.18	19.70	16.67	19.70	21.21	95.45	30.30	46.97
Pub. Transp	13.64	7.58	9.09	18.18	9.09	9.09	95.45	33.33	51.52
Comm.	7.58	12.12	7.58	7.58	9.09	4.55	100.00	22.73	56.06
Sound	23.64	14.55	20.00	27.27	30.91	30.91	96.36	27.27	69.09
Leisure	13.64	7.58	10.61	12.12	7.58	10.61	98.48	33.33	68.18
Books	22.73	15.15	19.70	18.18	16.67	16.67	95.45	18.18	42.42
Hotels	18.18	7.58	4.55	19.70	10.61	9.09	100.00	40.91	69.70
Average	21.40	13.30	15.93	21.21	15.17	17.37	96.49	26.79	55.41

Notes: ADF is the Augmented Dicky-Fuller unit root test, ADF-GLS is Elliot et al. (1996) test, ADF-WS is Park and Fuller's (1995) weighted symmetric test. Unit root tests are conducted at 10% significance level for $N_c(N_c + 1)/2$ distinct pairs of q_{ijct} , $i \neq j$, for each commodity c , and augmentation orders are chosen by AIC and SBC of ADF regression from augmentation orders $p = 0, 1, \dots, 12$, then the fraction of the rejected pairs over $N_c(N_c + 1)/2$ are computed. Critical values are different for T_{ijc} and augmentation orders p . The window to compute KPSS statistics is shown in the second row, the final column av uses a window of $(0.75T^{1/3} + 2T^{1/2})/2$. The 'Average' located at the bottom row of the table is a simple average of fractions over all the commodities (except CPI).

Table 4: ADF Unit Root Test Results for \bar{q}_t and \bar{q}_{ct}

	T_c	With an Intercept			With an Intercept and a Trend		
		ADF	10%CV	s_c	ADF	10%CV	s_c
\bar{q}_t	180	-1.498	-2.578	2	-1.678	-3.125	2
\bar{q}_{ct}							
Bread	180	-1.379	-2.578	2	-1.115	-3.125	2
Meat	180	-1.552	-2.578	2	-1.424	-3.125	2
Dairy	180	-1.295	-2.603	1	-1.019	-3.144	1
Fruits	180	-2.110	-2.559	12	-2.156	-3.082	12
Tobacco	180	-0.919	-2.603	1	-2.997	-3.144	1
Alcohol	180	-1.045	-2.572	4	-1.456	-3.122	4
Clothing	180	-1.352	-2.603	1	-1.243	-3.144	1
Footwear	180	-1.335	-2.603	1	-1.233	-3.144	1
Rents	180	-1.193	-2.603	1	-1.606	-3.144	1
Fuel	180	-1.998	-2.583	3	-2.520	-3.126	3
Furnit.	166	0.133	-2.586	1	-2.178	-3.144	1
Dom. Appl.	166	0.117	-2.586	1	-2.029	-3.144	1
Vehicles	180	-1.411	-2.578	2	-1.424	-3.125	2
Pub. Transp	180	-1.595	-2.578	2	-1.720	-3.125	2
Comm.	180	-2.392	-2.578	2	-2.901	-3.125	2
Sound	180	-1.527	-2.578	2	-0.626	-3.125	2
Leisure	177	-1.537	-2.578	2	-1.617	-3.125	2
Books	177	-1.320	-2.578	2	-0.851	-3.125	2
Hotels	177	-2.074	-2.559	12	-1.971	-3.082	12

Notes: The third and sixth columns of the table report ADF statistics, which are the t-ratios of $\hat{\phi}_c$ in the ADF regressions $\Delta\bar{q}_{ct} = \hat{\mu}_c + \hat{\delta}_{ct} + \hat{\phi}_c\bar{q}_{c,t-1} + \sum_{\ell=1}^{s_c} \hat{\xi}_{c\ell}\Delta\bar{q}_{c,t-\ell} + \hat{e}_{ct}$, $c = 1, 2, \dots, C$, $t = 14, 15, \dots, T_c$, where the lag augmentation s_c is chosen by AIC from $\ell = 1, 2, \dots, 12$. The ADF 10% critical values, which depend on T_c and s_c , are obtained via stochastic simulations with 10,000 replications.

Table 5: Average Fractions of Pairs for Which the Null is Rejected, Based on 2000 Replications of Bootstrapped Data. Regressions With an Intercept and Unit Roots not Imposed on the Commodity Factors

	Lag Order Chosen by AIC			Lag Order Chosen by SBC			KPSS		
	ADF	ADF-GLS	ADF-WS	ADF	ADF-GLS	ADF-WS	$0.75T^{1/3}$	$2T^{1/2}$	av
Subcategories									
Bread	37.79	43.16	49.23	30.86	37.52	41.36	74.02	20.19	34.47
Meat	48.56	62.51	68.64	40.22	56.71	61.33	67.39	13.79	27.03
Dairy	32.56	47.97	49.27	25.39	42.38	42.09	76.86	21.36	36.39
Fruits	64.03	64.07	76.48	82.73	82.05	90.05	36.96	8.16	15.92
Tobacco	47.65	44.90	56.23	43.57	42.62	52.67	72.52	18.53	32.11
Alcohol	26.06	27.42	32.22	23.35	24.47	28.29	61.13	23.73	34.73
Clothing	29.21	39.25	41.03	28.81	39.72	41.77	82.04	29.09	44.71
Footwear	36.05	44.85	46.78	28.32	40.39	40.99	76.81	27.67	41.23
Rents	29.66	39.22	42.61	20.23	31.60	33.14	76.15	18.92	33.99
Fuel	45.61	42.84	52.08	41.48	40.40	48.35	58.08	20.37	29.79
Furnit.	37.41	36.05	39.90	41.18	39.34	43.41	80.27	42.43	55.67
Dom. Appl.	37.75	39.57	44.65	42.05	43.48	48.13	79.51	37.99	51.76
Vehicles	23.98	34.20	35.50	17.18	26.65	26.07	82.14	25.36	41.95
Pub. Transp	29.75	34.22	39.00	26.53	31.19	34.94	80.61	27.40	42.55
Comm.	34.35	24.78	35.65	28.96	22.78	30.67	79.36	28.57	42.89
Sound	59.03	58.63	70.37	62.18	62.23	72.86	61.56	15.51	26.33
Leisure	21.92	37.08	42.16	23.97	36.76	43.29	67.47	21.03	34.68
Books	29.82	41.77	43.24	22.41	35.38	35.33	78.61	24.37	39.65
Hotels	53.76	67.50	70.84	58.25	71.82	75.55	63.46	15.22	26.54
Average	38.16	43.68	49.26	36.19	42.50	46.86	71.31	23.14	36.44

Notes: This table reports the average fractions, $R^{-1} \sum_{r=1}^R F_c^{(r)}$, $c = 1, 2, \dots, C$, where $F_c^{(r)}$ is the fraction of the pairs, $q_{ijct}^{(r)} = q_{ict}^{(r)} - q_{jct}^{(r)}$, in the r^{th} replication for which the null hypothesis is rejected by the tests. The r^{th} replication of q_{ict} is generated as $q_{ict}^{(r)} = \hat{\alpha}_{ic} + \hat{\gamma}_{ic1} \bar{q}_{ct}^{(r)} + \hat{\gamma}_{ic2} \bar{q}_t^{(r)} + \varepsilon_{ict}^{(r)}$, $i = 1, 2, \dots, N_c$; $t = 1, 2, \dots, T$; $r = 1, 2, \dots, R$, where the estimates $\hat{\alpha}_{ic}$, $\hat{\gamma}_{ic1}$, $\hat{\gamma}_{ic2}$, $\hat{\varepsilon}_{ict}$ are obtained from the regression of q_{ict} on unity, \bar{q}_{ct} and \bar{q}_t for each combination of (i, c) , $\varepsilon_{ict}^{(r)}$ are generated based on the estimates of ADF regression of $\hat{\varepsilon}_{ict}$ and random draws of its residuals with replacement. $\bar{q}_{ct}^{(r)}$ is generated based on the estimates of ADF regression of \bar{q}_{ct} and random draws of its residuals with replacement. $\bar{q}_t^{(r)} \equiv C^{-1} \sum_{c=1}^C \bar{q}_{ct}^{(r)}$. $T = 180$, and $R = 2000$. See Section 4 for the details of bootstrap procedure. See also notes to Table 3.

Table 6: Average Fractions of Pairs for Which the Null is Rejected, Based on 2000 Replications of Bootstrapped Data. Regressions With an Intercept and Unit Roots Imposed on the Commodity Factors

	Lag Order Chosen by AIC			Lag Order Chosen by SBC			KPSS		
	ADF	ADF-GLS	ADF-WS	ADF	ADF-GLS	ADF-WS	$0.75T^{1/3}$	$2T^{1/2}$	av
Subcategories									
Bread	27.41	27.19	31.24	22.94	24.09	26.86	83.95	39.77	54.10
Meat	31.10	30.07	36.72	26.66	27.01	32.65	82.43	40.55	53.90
Dairy	23.03	27.26	28.78	18.47	23.84	24.28	86.41	42.57	57.29
Fruits	37.88	41.56	44.15	53.84	54.99	58.63	68.14	35.63	48.33
Tobacco	41.10	38.20	46.57	37.76	36.42	43.63	78.74	30.41	44.22
Alcohol	21.37	19.96	23.84	19.65	18.44	21.33	65.53	34.49	44.62
Clothing	22.65	25.79	27.63	22.57	26.19	28.29	88.12	43.37	58.40
Footwear	28.78	32.75	35.22	22.54	29.01	30.32	83.90	41.03	54.23
Rents	23.14	23.53	26.90	15.85	18.41	20.23	84.04	38.93	52.43
Fuel	26.98	27.94	30.63	25.03	26.41	28.33	70.90	37.73	48.30
Furnit.	37.15	35.72	39.65	40.96	39.05	43.23	80.48	42.81	55.78
Dom. Appl.	36.15	37.41	42.24	40.56	41.33	45.98	81.03	39.90	53.41
Vehicles	17.95	18.27	20.35	13.40	14.05	14.68	88.59	42.20	57.24
Pub. Transp	19.85	19.43	21.68	18.68	18.16	20.10	89.08	44.21	59.37
Comm.	20.29	24.99	26.84	17.35	22.73	23.80	86.98	41.12	56.39
Sound	40.90	35.65	44.60	44.57	39.28	47.90	78.07	37.08	49.84
Leisure	13.21	14.79	20.51	13.96	13.88	22.84	75.88	40.60	52.69
Books	21.91	24.88	26.94	16.68	20.49	21.42	86.84	42.67	56.88
Hotels	23.68	22.09	26.31	26.00	24.52	29.12	87.40	44.98	59.13
Average	27.08	27.76	31.62	26.18	27.28	30.72	81.40	40.00	53.50

Notes: See notes to Tables 3 and 5. The simulation design is identical to that of Table 5, except $\bar{q}_{ct}^{(r)}$ is restricted to have a unit root without drift. Namely, (14) is replaced with $\bar{q}_{ct}^{(r)} = \bar{q}_{c,t-1}^{(r)} + \sum_{\ell=1}^{s_c} \hat{\xi}_{c\ell} \Delta \bar{q}_{c,t-\ell}^{(r)} + e_{ct}^{(r)}$, $t = 1, 2, \dots, T$; $c = 1, 2, \dots, C$, $r = 1, 2, \dots, R$, $R = 2000$.

SUPPLEMENT

for

Pairwise Tests of Purchasing Power Parity Using Aggregate and Disaggregate Price Measures

by

M. Hashem Pesaran, Ron P. Smith, Takashi Yamagata,
Liudmyla Hvozdyk

March 30, 2006

- 1: Real Effective Exchange Rates: Procedures and Results
- 2: Unit Root Tests With Just Intercept
- 3: Seasonal Adjustment: Procedures and Results
- 4: Critical Values
- 5: Average Lags
- 6: Estimates of the Commodity Factor Models
- 7: Countries With Similar Factor Loadings

1 Real Effective Exchange Rates: Procedures and Results

Define M_{ijt} and X_{ijt} to be import of i from j and export from i to j at time t , respectively. The total trade between i and j is defined as

$$D_{ijt} = M_{ijt} + X_{ijt}.$$

We compute fixed trade weights by averaging trade flows over a three year time interval 1987–1989, using the data on imports and exports from the IMF Direction of Trade Statistics (DOTS) as

$$w_{i,j} = \frac{\bar{D}_{i,j}}{\sum_{j=0}^N \bar{D}_{i,j}}$$

where $\bar{D}_{i,j} = \sum_{t=1987}^{1989} D_{i,j}/3$. Observe that $w_{ii} = 0$ and $\sum_{j=0}^N w_{ij} = 1$. The weights are shown as below:

Table S1.1: Trade Weights

(i,j)	BE	DK	FI	FR	DE	GR	IT	NL	PT	ES	UK	US
BE	0	0.0100	0.0067	0.2287	0.2788	0.0049	0.0680	0.2046	0.0061	0.0231	0.1083	0.0608
DK	0.0456	0	0.0473	0.0926	0.3431	0.0094	0.0740	0.0838	0.0133	0.0239	0.1636	0.1034
FI	0.0463	0.0662	0	0.0971	0.2853	0.0087	0.0740	0.0704	0.0131	0.0272	0.1924	0.1193
FR	0.1343	0.0123	0.0093	0	0.2650	0.0085	0.1743	0.0794	0.0163	0.0723	0.1210	0.1072
DE	0.1175	0.0317	0.0179	0.2009	0	0.0142	0.1491	0.1544	0.0125	0.0439	0.1335	0.1243
GR	0.0483	0.0175	0.0120	0.1156	0.3163	0	0.2222	0.0870	0.0042	0.0246	0.0894	0.0630
IT	0.0651	0.0139	0.0090	0.2423	0.3066	0.0205	0	0.0685	0.0116	0.0515	0.0994	0.1116
NL	0.1920	0.0187	0.0114	0.1214	0.3478	0.0079	0.0672	0	0.0080	0.0231	0.1225	0.0800
PT	0.0472	0.0192	0.0126	0.1766	0.1989	0.0035	0.0958	0.0712	0	0.1702	0.1355	0.0692
ES	0.0470	0.0104	0.0081	0.2283	0.2125	0.0064	0.1385	0.0579	0.0503	0	0.1174	0.1231
UK	0.0804	0.0280	0.0224	0.1555	0.2397	0.0074	0.0878	0.1237	0.0151	0.0457	0	0.1944
US	0.0751	0.0162	0.0127	0.1398	0.2582	0.0071	0.1169	0.0914	0.0095	0.0465	0.2267	0

Notes: The source of the data is http://www.esds.ac.uk/international/support/user_guides/imf/dots.asp

Table S1.2: ADF Test Results of Real Effective Exchange Rate, $\bar{q}_{it} = \sum_{j=0}^N w_{ij}q_{ijt}$, With an Intercept

	Test	DF	ADF(1)	ADF(2)	ADF(3)	ADF(4)	ADF(5)	ADF(6)	ADF(7)	ADF(8)	ADF(9)	ADF(10)	ADF(11)	ADF(12)
BE	Statistic	-1.013	-1.602	-1.548	-1.825	-1.705	-1.679	-1.323	-1.489	-1.479	-1.616	-1.497	-1.649	-1.399
	AIC	570.013	575.964	574.977	575.672	574.781	573.784	574.205	573.732	572.740	572.152	571.350	570.971	571.254
DE	Statistic	-0.419	-0.777	-0.863	-0.840	-0.676	-0.697	-0.904	-1.040	-0.966	-1.097	-1.390	-1.519	-1.728
	AIC	555.388	555.388	554.539	553.540	553.092	552.110	551.790	551.105	550.118	549.351	549.717	549.077	548.856
DK	Statistic	-0.823	-1.232	-1.100	-1.325	-1.229	-1.218	-1.086	-0.953	-1.093	-1.066	-0.994	-0.916	-0.940
	AIC	568.716	574.972	574.496	575.657	574.857	573.857	573.340	572.936	572.762	571.781	571.050	570.492	569.556
ES	Statistic	-0.910	-1.257	-1.254	-1.290	-1.153	-1.183	-1.358	-1.489	-1.392	-1.485	-1.403	-1.289	-1.241
	AIC	490.756	499.201	498.204	497.318	497.225	496.313	497.108	496.976	496.259	495.735	494.934	494.468	493.546
IT	Statistic	-1.072	-1.524	-1.208	-1.566	-1.266	-1.416	-1.397	-1.293	-1.298	-1.386	-1.468	-1.459	-1.588
	AIC	464.282	471.204	473.074	477.193	478.203	477.875	476.875	476.001	475.020	474.248	473.557	472.559	472.175
FR	Statistic	-1.780	-2.528	-2.221	-2.487	-2.127	-1.988	-1.681	-1.990	-2.270	-2.317	-2.329	-2.187	-1.960
	AIC	552.808	558.853	558.590	558.835	559.262	558.361	558.790	559.774	559.911	559.043	558.095	557.246	557.133
GR	Statistic	-1.626	-1.910	-1.600	-1.444	-1.287	-1.242	-1.100	-1.320	-1.313	-1.074	-0.877	-0.807	-0.925
	AIC	435.685	436.879	437.622	437.081	436.625	435.647	435.105	435.565	434.568	434.899	435.037	434.181	433.773
NL	Statistic	-1.112	-1.520	-1.598	-1.392	-1.487	-1.564	-1.933	-1.877	-1.808	-2.257	-2.187	-2.152	-2.784
	AIC	602.563	604.781	603.953	603.530	602.792	601.965	603.001	602.001	601.003	602.179	601.179	600.185	604.073
PT	Statistic	-0.358	-0.774	-0.549	-0.676	-0.580	-0.418	-0.298	-0.260	-0.314	-0.444	-0.456	-0.454	-0.612
	AIC	480.016	484.189	484.593	484.149	483.373	483.105	482.526	481.560	480.743	480.324	479.335	478.335	478.324
FI	Statistic	-0.981	-1.320	-1.295	-1.278	-1.365	-1.584	-1.718	-1.494	-1.524	-1.667	-1.937	-1.804	-2.120
	AIC	455.064	460.633	459.640	458.641	458.051	458.837	458.461	458.606	457.678	457.417	458.266	457.461	458.689
UK	Statistic	-1.759	-2.641	-2.515	-2.667	-2.657	-2.151	-2.151	-2.220	-2.162	-2.172	-2.005	-2.074	-2.039
	AIC	429.512	441.896	440.939	440.448	439.490	441.741	440.764	440.014	439.040	438.078	437.824	437.303	436.319
US	Statistic	-0.809	-1.245	-1.139	-1.308	-1.247	-1.215	-1.164	-1.201	-1.317	-1.424	-1.498	-1.435	-1.554
	AIC	368.561	375.657	374.962	375.153	374.227	373.239	372.287	371.371	370.888	370.334	369.585	368.645	368.243
CV	10%	-2.586	-2.584	-2.574	-2.572	-2.579	-2.562	-2.549	-2.541	-2.535	-2.537	-2.525	-2.547	-2.533
CV	5%	-2.882	-2.880	-2.872	-2.884	-2.872	-2.872	-2.864	-2.884	-2.864	-2.872	-2.840	-2.867	-2.854

Notes: The real effective exchange rate, $\bar{q}_{it} = \sum_{j=0}^N w_{ij}q_{ijt}$, is based on the CPI of the countries. The lag augmentation is chosen by AIC from $\ell = 1, 2, \dots, 12$. In bold is the maximum AIC value and the corresponding test statistic. The last two rows represent critical values which depend upon T and are obtained by simulation with 10000 replications.

Table S1.3: ADF Test Results of Real Effective Exchange Rate, $\bar{q}_{it} = \sum_{j=0}^N w_{ij}q_{ijt}$, With an Intercept and a Trend

	Test	DF	ADF(1)	ADF(2)	ADF(3)	ADF(4)	ADF(5)	ADF(6)	ADF(7)	ADF(8)	ADF(9)	ADF(10)	ADF(11)	ADF(12)
BE	Statistic	-4.579	-5.090	-5.046	-5.382	-5.279	-5.306	-4.969	-5.137	-5.127	-5.238	-5.139	-5.245	-5.101
	AIC	579.588	586.541	585.559	587.255	586.255	585.539	585.321	585.338	584.436	584.136	583.280	583.106	583.532
DE	Statistic	-1.378	-1.716	-1.794	-1.773	-1.594	-1.621	-1.838	-1.989	-1.925	-2.060	-2.390	-2.574	-2.858
	AIC	554.003	555.826	555.027	554.028	553.440	552.493	552.345	551.803	550.803	550.160	550.865	550.506	550.723
DK	Statistic	-1.707	-2.275	-2.106	-2.465	-2.353	-2.361	-2.193	-2.018	-2.229	-2.203	-2.090	-1.950	-2.006
	AIC	568.860	575.814	575.140	576.852	575.921	574.966	574.235	573.609	573.754	572.758	571.863	571.094	570.263
ES	Statistic	-1.114	-1.602	-1.606	-1.668	-1.494	-1.536	-1.778	-1.971	-1.843	-2.018	-1.905	-1.718	-1.655
	AIC	489.971	498.706	497.721	496.894	496.695	495.817	496.800	496.852	496.033	495.728	494.826	494.170	493.204
IT	Statistic	-1.161	-1.557	-1.209	-1.535	-1.193	-1.297	-1.214	-1.016	-0.836	-0.837	-0.897	-0.837	-0.963
	AIC	465.902	471.609	474.283	477.531	479.089	478.407	477.462	476.899	476.022	475.029	474.128	473.161	472.481
FR	Statistic	-2.600	-3.267	-2.972	-3.195	-2.880	-2.739	-2.473	-2.718	-2.936	-2.952	-2.943	-2.798	-2.594
	AIC	554.331	560.373	560.083	560.289	560.891	560.016	560.683	561.508	561.518	560.598	559.627	558.829	558.949
GR	Statistic	-2.189	-2.402	-2.125	-1.979	-1.834	-1.774	-1.636	-1.785	-1.753	-1.525	-1.335	-1.247	-1.309
	AIC	436.746	437.686	438.662	438.298	438.076	437.156	436.882	436.988	435.993	436.693	437.253	436.633	435.921
NL	Statistic	-1.253	-1.628	-1.692	-1.486	-1.578	-1.652	-2.013	-1.953	-1.877	-2.307	-2.232	-2.193	-2.811
	AIC	602.203	604.275	603.417	603.028	602.290	601.462	602.473	601.474	600.476	601.570	600.570	599.575	603.380
PT	Statistic	-2.600	-2.843	-2.657	-2.742	-2.663	-2.558	-2.475	-2.440	-2.483	-2.592	-2.607	-2.602	-2.818
	AIC	483.214	487.432	487.684	487.313	486.492	486.235	485.662	484.696	483.948	483.677	482.741	481.759	482.223
FI	Statistic	-1.188	-1.501	-1.476	-1.458	-1.541	-1.750	-1.878	-1.661	-1.687	-1.816	-2.063	-1.936	-2.227
	AIC	454.438	459.982	458.990	457.991	457.402	458.193	457.820	457.962	457.032	456.753	457.567	456.774	457.966
UK	Statistic	-1.943	-2.914	-2.789	-2.966	-2.970	-2.404	-2.414	-2.510	-2.448	-2.472	-2.264	-2.380	-2.340
	AIC	428.860	441.659	440.675	440.302	439.389	441.342	440.393	439.733	438.738	437.819	437.423	437.038	436.038
US	Statistic	-1.988	-2.414	-2.306	-2.518	-2.453	-2.425	-2.374	-2.420	-2.536	-2.636	-2.720	-2.658	-2.765
	AIC	369.442	376.910	376.125	376.606	375.620	374.622	373.631	372.789	372.438	371.992	371.365	370.381	370.085
CV	10%	-3.138	-3.124	-3.124	-3.131	-3.127	-3.126	-3.110	-3.100	-3.101	-3.080	-3.066	-3.080	-3.058
CV	5%	-3.447	-3.445	-3.427	-3.424	-3.414	-3.419	-3.404	-3.413	-3.412	-3.429	-3.393	-3.388	-3.370

Notes: The real effective exchange rate, $\bar{q}_{it} = \sum_{j=0}^N w_{ij}q_{ijt}$, is based on the CPI of the countries. The lag augmentation is chosen by AIC from $\ell = 1, 2, \dots, 12$. In bold is the maximum AIC value and the corresponding test statistic. The last two rows represent critical values which depend upon T and are obtained by simulation with 10000 replications.

2 Unit Root Tests With Just Intercept

Table S2.1: Fractions of Pairs q_{ijct} for Which the Unit Root Hypothesis is Rejected at 10% Significance Level in the Case With an Intercept, for All 12 Countries

	Lag Order Chosen by AIC			Lag Order Chosen by SBC			KPSS		
	ADF	ADF-GLS	ADF-WS	ADF	ADF-GLS	ADF-WS	$0.75T^{1/3}$	$2T^{1/2}$	av
CPI	7.58	6.06	7.58	9.09	10.61	10.61	96.97	31.82	53.03
Subcategories									
Bread	6.06	4.55	6.06	9.09	7.58	10.61	96.97	50.00	72.73
Meat	12.12	19.70	19.70	10.61	18.18	19.70	95.45	25.76	50.00
Dairy	22.73	13.64	21.21	24.24	12.12	19.70	89.39	28.79	51.52
Fruits	28.79	18.18	30.30	53.03	46.97	65.15	77.27	37.88	50.00
Tobacco	21.21	19.70	21.21	21.21	16.67	22.73	81.82	28.79	46.97
Alcohol	9.09	18.18	19.70	6.06	18.18	16.67	93.94	33.33	59.09
Clothing	6.06	10.61	10.61	6.06	9.09	9.09	92.42	60.61	77.27
Footwear	18.18	6.06	13.64	15.15	6.06	7.58	86.36	37.88	46.97
Rents	18.18	9.09	12.73	14.55	5.45	7.27	94.55	30.91	49.09
Fuel	21.82	18.18	27.27	38.18	23.64	34.55	90.91	23.64	49.09
Furnit.	12.73	5.45	5.45	10.91	3.64	3.64	90.91	50.91	67.27
Dom. Appl.	9.09	9.09	5.45	5.45	5.45	5.45	94.55	56.36	67.27
Vehicles	7.58	3.03	6.06	10.61	3.03	4.55	95.45	45.45	63.64
Pub. Transp	13.64	9.09	12.12	19.70	7.58	12.12	90.91	24.24	46.97
Comm.	6.06	15.15	15.15	6.06	16.67	18.18	90.91	33.33	51.52
Sound	9.09	12.73	14.55	18.18	14.55	20.00	87.27	36.36	47.27
Leisure	16.67	9.09	13.64	15.15	7.58	9.09	98.48	22.73	57.58
Books	10.61	10.61	7.58	9.09	9.09	9.09	93.94	39.39	69.70
Hotels	7.58	10.61	16.67	4.55	7.58	13.64	96.97	33.33	60.61
Average	13.54	11.72	14.69	15.68	12.58	16.25	91.50	36.83	57.08

Notes: ADF is a standard Dicky-Fuller unit root test, ADF-GLS is Elliot et al. (1996) test, ADF-WS is Park and Fuller's (1995) weighted symmetric test. Unit root tests are conducted at 10% significance level for $N_c(N_c + 1)/2$ distinct pairs of q_{ijct} , $i \neq j$, for each commodity c , and augmentation orders are chosen by AIC and SBC of ADF regression from augmentation orders $p = 0, 1, \dots, 12$, then the fraction of the rejected pairs over $N_c(N_c + 1)/2$ are computed. Critical values are tabulated in Section 4, which are different for T_{ijc} and augmentation orders. Lag window to compute KPSS statistics is shown in the second row. The category 'Average' located at the bottom row of the table is a simple average of fractions of all the commodities except for CPI.

3 Seasonal Adjustment: Procedures and Results

Price data series can be written as (suppressing i, c)

$$P_{ym}, y = 1, 2, \dots, Y_m; m = 1, 2, \dots, 12,$$

where y indicates year, $y = 1, 2, \dots, Y_m$, and m stands for month, $m = 1, 2, \dots, 12$. We use, as usually, lower case to indicate logarithm:

$$p_{ym} = \ln(P_{ym}).$$

Next, we take the first difference:

$$\Delta p_{ym} = p_{ym} - p_{y,m-1}.$$

Note that for p_{ym} with $m = 1$, $p_{y,m-1} \equiv p_{y-1,12}$. As the initial observation is lost, it is replaced with $\Delta p_{y+1,1}$, which is the lagged price of the initial month of the second year. Then, take the monthly average of Δp_{ym} :

$$\Delta \bar{p}_m = \frac{1}{Y_m} \sum_{y=1}^{Y_m} \Delta p_{ym}.$$

Note that the number of years, Y_m , depends on the month. Also, as $\sum_{m=1}^M \Delta \bar{p}_m \neq 0$ we compute demeaned monthly average, $\Delta \bar{p}_m^+ = \Delta \bar{p}_m - a$, $a = M^{-1} \sum_{m=1}^M \Delta \bar{p}_m$, and define monthly demeaned lagged price as

$$\Delta \hat{p}_{ym} = \Delta p_{ym} - \Delta \bar{p}_m^+.$$

This ensures that

$$\sum_{m=1}^M \Delta \hat{p}_{ym} = \sum_{m=1}^M \Delta p_{ym} \text{ for all } y,$$

which we wanted. Next, we recover level price by taking a cumulative sum. Here, the initial observation is replaced with the original initial observations of p_{ym} , and we find seasonally adjusted log of price, p_{ym}^* . Transform double index to the single index, i.e. $\{p_{11}^*, p_{12}^*, \dots, p_{1m}^*; p_{21}^*, p_{22}^*, \dots, p_{2m}^*; \dots; p_{y1}^*, p_{y2}^*, \dots, p_{ym}^*\}$ to p_t^* , $t = 1, 2, \dots, T$, $T = 12Y_m$. We define the log of seasonally adjusted price index of country i for commodity c at time t as

$$q_{ict}^* = e_{it} - p_{ct}^* + p_{ict}^*,$$

where e_{it} is the log of the bilateral exchange rate between country i and the US, p_{ct}^* is the log of seasonally adjusted US price index, and p_{ict}^* is the log of the seasonally adjusted price index of country i 's commodity c at time t . The log of seasonally adjusted real exchange rate for commodity c between any other pair of countries $i, j \neq 0$ is computed as

$$q_{ijct}^* = q_{ict}^* - q_{jct}^*.$$

Criterion for Seasonal Adjustment Application

We apply seasonal adjustment described above if the following criteria are satisfied. First, we run a joint significance test of the regression of Δp_{ict} on seasonal dummies for all commodity categories across countries at 1% level. Next, if the joint tests are significant for a half of the countries or more, we examine whether the time series plot of p_{ict} exhibits seasonality pattern. Otherwise, no seasonal adjustment is applied. We also do not apply seasonal adjustment if the level price exhibits a step function. The summary of seasonal adjustment is reported below.

Table S3.1: Summary on Seasonal Adjustment Application

Commodity	# rej at 1%	Description	Adjust?
Bread	5/12	Step function	No
Meat	8/12	Not step function except DK	Yes except BE, DE,DK,FE,US
Dairy	4/12	Only DE, ES, NL, US's dummies are jointly significant. For DE and ES seasonal adjustment seems effective	No
Fruits	12/12	It may be reasonable to seasonally adjust all	Yes
Tobacco	6/12	Clear step function	No
Alcohol	5/12	DK and FI are step functions, but others are smooth.	No
Clothing	11/12	Except US, seasonality exists (For us, p-value is 0.88)	Yes, except US
Footwear	9/12	Except US, seasonality exists (For us, p-value is 0.999)	Yes, except US
Rents	10/12	Most series are step functions	No
Fuel	4/12	No seasonality	No
Furnit.	10/12	It may be reasonable to seasonally adjust all	Yes
Dom. Appl.	7/12	It may be reasonable to seasonally adjust all	Yes
Vehicles	6/12	It is not very clear. DK is step function.	Yes except DK
Pub. Transp	11/12	Clear step function is only GR, others are kind of mixture.	No
Comm.	5/12	Most series are step functions	No
Sound	8/11	Sudden jump for US after 1995M10. DK is a sort of step function	Yes except DK
Leisure	11/12	Mixture of step type movement and seasonality.	No
Books	4/12	Step function	No
Hotels	11/12	DK and FI are step functions. Higher seasonality for BE after 1992M1, for NL after 1991M1.	Yes, except DK, FI,BE,NL
CPI	11/12	All to be adjusted	Yes

Table S3.2: Fractions of Pairs q_{ijct}^* (Seasonally Adjusted) for Which the Unit Root Hypothesis is Rejected at 10% Significance Level in the Case with an Intercept and a Trend, for All 12 Countries

	Lag Order Chosen by AIC			Lag Order Chosen by SBC			KPSS	
	ADF	ADF-GLS	ADF-WS	ADF	ADF-GLS	ADF-WS	$0.75T^{1/3}$	$2T^{1/2}$
CPI	13.64	4.55	9.09	12.12	10.61	12.12	96.97	31.82
Subcategories								
Bread	13.64	13.64	12.12	13.64	7.58	10.61	98.48	22.73
Meat	21.21	10.61	15.15	13.64	10.61	15.15	98.48	28.79
Dairy	13.64	13.64	12.12	13.64	7.58	10.61	98.48	22.73
Fruits	42.42	39.39	39.39	57.58	54.55	59.09	93.94	24.24
Tobacco	40.91	15.15	22.73	42.42	19.70	25.76	93.94	16.67
Alcohol	19.70	16.67	18.18	15.15	12.12	13.64	96.97	28.79
Clothing	12.12	7.58	7.58	18.18	3.03	4.55	98.48	18.18
Footwear	22.73	1.52	6.06	21.21	3.03	6.06	96.97	28.79
Rents	16.36	7.27	10.91	12.73	1.82	3.64	98.18	34.55
Fuel	27.27	23.64	27.27	32.73	30.91	34.55	96.36	30.91
Furnit.	27.27	9.09	16.67	24.24	6.06	12.12	98.48	21.21
Dom. Appl.	27.27	15.15	21.21	15.15	7.58	12.12	93.94	12.12
Vehicles	18.18	15.15	19.70	15.15	13.64	21.21	95.45	30.30
Pub. Transp	13.64	7.58	9.09	18.18	9.09	9.09	95.45	33.33
Comm.	7.58	12.12	7.58	7.58	9.09	4.55	100.00	22.73
Sound	25.45	16.36	21.82	25.45	20.00	21.82	98.18	27.27
Leisure	13.64	7.58	10.61	12.12	7.58	10.61	98.48	33.33
Books	22.73	15.15	19.70	18.18	16.67	16.67	95.45	18.18
Hotels	10.61	12.12	9.09	12.12	7.58	7.58	100.00	40.91
Average	20.86	13.65	16.16	20.48	13.06	15.76	97.14	26.09

Notes: ADF is a standard Dicky-Fuller unit root test, ADF-GLS is Elliot et al. (1996) test, ADF-WS is Park and Fuller's (1995) weighted symmetric test. Unit root tests are conducted at 10% significance level for $N_c(N_c + 1)/2$ distinct pairs of q_{ijct}^* , $i \neq j$, where q_{ijct}^* is based on seasonally adjusted prices, for each commodity c , and augmentation orders are chosen by AIC and SBC of ADF regression from augmentation orders $p = 0, 1, \dots, 12$, then the fraction of the rejected pairs over $N_c(N_c + 1)/2$ are computed. Critical values are tabulated in Section 4, which are different for T_{ijc} and augmentation orders. Lag window to compute KPSS statistics is shown at the second row. The category 'Average' located at the bottom row of the table is a simple average of fractions of all the commodities except CPI.

Table S3.3: Fractions of Pairs q_{ijct}^* (Seasonally Adjusted) for Which the Unit Root Hypothesis is Rejected at 10% Significance Level in the Case With an Intercept, for All 12 Countries

	Lag Order Chosen by AIC			Lag Order Chosen by SBC			KPSS		
	ADF	ADF-GLS	ADF-WS	ADF	ADF-GLS	ADF-WS	$0.75T^{1/3}$	$2T^{1/2}$	av
CPI	7.58	9.09	10.61	7.58	10.61	7.58	96.97	31.82	53.03
Subcategories									
Bread	6.06	4.55	6.06	9.09	7.58	10.61	96.97	50.00	72.73
Meat	13.64	10.61	16.67	10.61	12.12	15.15	95.45	25.76	50.00
Dairy	6.06	4.55	6.06	9.09	7.58	10.61	96.97	50.00	72.73
Fruits	36.36	30.30	39.39	57.58	45.45	62.12	84.85	36.36	51.52
Tobacco	21.21	19.70	21.21	21.21	16.67	22.73	81.82	28.79	46.97
Alcohol	9.09	18.18	19.70	6.06	18.18	16.67	93.94	33.33	59.09
Clothing	6.06	6.06	3.03	4.55	1.52	0.00	92.42	60.61	77.27
Footwear	13.64	4.55	6.06	12.12	4.55	3.03	87.88	37.88	46.97
Rents	18.18	9.09	12.73	14.55	5.45	7.27	94.55	30.91	49.09
Fuel	21.82	18.18	27.27	38.18	23.64	34.55	90.91	23.64	49.09
Furnit.	18.18	6.06	10.61	16.67	7.58	10.61	92.42	57.58	71.21
Dom. Appl.	9.09	12.12	10.61	7.58	13.64	12.12	95.45	62.12	71.21
Vehicles	6.06	3.03	4.55	10.61	3.03	4.55	95.45	45.45	63.64
Pub. Transp	13.64	9.09	12.12	19.70	7.58	12.12	90.91	24.24	46.97
Comm.	6.06	15.15	15.15	6.06	16.67	18.18	90.91	33.33	51.52
Sound	12.73	10.91	14.55	12.73	10.91	14.55	87.27	36.36	47.27
Leisure	16.67	9.09	13.64	15.15	7.58	9.09	98.48	22.73	57.58
Books	10.61	10.61	7.58	9.09	9.09	9.09	93.94	39.39	69.70
Hotels	4.55	10.61	10.61	1.52	4.55	4.55	96.97	33.33	60.61
Average	13.14	11.18	13.56	14.85	11.76	14.61	92.50	38.52	58.69

Notes: ADF is a standard Dicky-Fuller unit root test, ADF-GLS is Elliot et al. (1996) test, ADF-WS is Park and Fuller's (1995) weighted symmetric test. Unit root tests are conducted at 10% significance level for $N_c(N_c + 1)/2$ distinct pairs of q_{ijct}^* , $i \neq j$, where q_{ijct}^* is based on seasonally adjusted prices, for each commodity c , and augmentation orders are chosen by AIC and SBC of ADF regression from augmentation orders $p = 0, 1, \dots, 12$, then the fraction of the rejected pairs over $N_c(N_c + 1)/2$ are computed. Critical values are tabulated in Section 4, which are different for T_{ijc} and augmentation orders. Lag window to compute KPSS statistics is shown at the second row. The category 'Average' located at the bottom row of the table is a simple average of fractions of all the commodities except CPI.

4 Critical Values

p,T	With Intercept, 10%					With Intercept and Trend, 10%				
	84	125	132	165-166	175-180	84	125	132	165-166	175-180
	ADF					ADF				
0	-2.594	-2.549	-2.590	-2.586	-2.601	-3.148	-3.151	-3.141	-3.138	-3.130
1	-2.604	-2.571	-2.600	-2.584	-2.603	-3.143	-3.141	-3.148	-3.124	-3.144
2	-2.563	-2.548	-2.564	-2.570	-2.578	-3.122	-3.116	-3.132	-3.124	-3.125
3	-2.592	-2.557	-2.584	-2.572	-2.583	-3.127	-3.124	-3.128	-3.131	-3.126
4	-2.557	-2.547	-2.572	-2.579	-2.572	-3.129	-3.117	-3.120	-3.127	-3.122
5	-2.587	-2.560	-2.576	-2.562	-2.576	-3.116	-3.121	-3.115	-3.126	-3.124
6	-2.534	-2.536	-2.569	-2.549	-2.570	-3.083	-3.098	-3.116	-3.110	-3.118
7	-2.531	-2.533	-2.551	-2.541	-2.576	-3.087	-3.089	-3.117	-3.100	-3.109
8	-2.519	-2.523	-2.539	-2.535	-2.569	-3.062	-3.071	-3.111	-3.101	-3.094
9	-2.529	-2.530	-2.540	-2.537	-2.583	-3.068	-3.075	-3.111	-3.080	-3.111
10	-2.504	-2.522	-2.531	-2.525	-2.574	-3.036	-3.053	-3.090	-3.066	-3.077
11	-2.495	-2.529	-2.526	-2.547	-2.561	-3.030	-3.070	-3.090	-3.080	-3.079
12	-2.491	-2.516	-2.522	-2.533	-2.559	-3.007	-3.042	-3.071	-3.058	-3.082
	ADF-GLS					ADF-GLS				
0	-1.876	-1.793	-1.797	-1.783	-1.754	-2.764	-2.691	-2.684	-2.656	-2.650
1	-1.866	-1.786	-1.789	-1.793	-1.744	-2.759	-2.695	-2.681	-2.663	-2.664
2	-1.843	-1.772	-1.772	-1.777	-1.740	-2.715	-2.667	-2.664	-2.642	-2.638
3	-1.843	-1.778	-1.787	-1.763	-1.728	-2.714	-2.675	-2.677	-2.651	-2.636
4	-1.830	-1.767	-1.773	-1.745	-1.736	-2.691	-2.651	-2.659	-2.638	-2.625
5	-1.833	-1.772	-1.781	-1.738	-1.733	-2.698	-2.666	-2.643	-2.643	-2.630
6	-1.799	-1.748	-1.760	-1.727	-1.722	-2.655	-2.635	-2.640	-2.635	-2.620
7	-1.799	-1.749	-1.751	-1.729	-1.727	-2.642	-2.637	-2.630	-2.625	-2.618
8	-1.789	-1.731	-1.743	-1.717	-1.707	-2.594	-2.619	-2.606	-2.606	-2.600
9	-1.785	-1.720	-1.748	-1.713	-1.711	-2.610	-2.622	-2.599	-2.605	-2.615
10	-1.768	-1.708	-1.725	-1.707	-1.697	-2.562	-2.600	-2.585	-2.576	-2.581
11	-1.750	-1.712	-1.728	-1.707	-1.693	-2.541	-2.594	-2.581	-2.582	-2.584
12	-1.728	-1.696	-1.710	-1.696	-1.674	-2.501	-2.555	-2.555	-2.558	-2.581
	ADF-WS					ADF-WS				
0	-2.279	-2.254	-2.272	-2.275	-2.256	-2.963	-2.932	-2.933	-2.928	-2.928
1	-2.285	-2.268	-2.279	-2.279	-2.274	-2.983	-2.954	-2.958	-2.965	-2.937
2	-2.279	-2.257	-2.267	-2.269	-2.260	-2.969	-2.951	-2.937	-2.939	-2.939
3	-2.297	-2.269	-2.282	-2.288	-2.269	-2.993	-2.976	-2.976	-2.952	-2.946
4	-2.278	-2.262	-2.286	-2.258	-2.261	-2.991	-2.964	-2.983	-2.966	-2.947
5	-2.313	-2.288	-2.287	-2.274	-2.277	-3.034	-2.990	-2.998	-2.982	-2.985
6	-2.306	-2.289	-2.276	-2.268	-2.281	-3.021	-2.980	-2.983	-2.968	-2.964
7	-2.327	-2.277	-2.271	-2.283	-2.285	-3.034	-3.006	-3.003	-2.981	-2.985
8	-2.318	-2.271	-2.278	-2.276	-2.268	-3.017	-2.999	-2.999	-2.980	-2.969
9	-2.329	-2.284	-2.286	-2.291	-2.278	-3.055	-3.017	-3.002	-2.998	-2.980
10	-2.315	-2.284	-2.289	-2.270	-2.264	-3.016	-3.006	-2.993	-2.964	-2.973
11	-2.331	-2.298	-2.302	-2.286	-2.279	-3.032	-3.016	-3.008	-2.983	-2.988
12	-2.313	-2.300	-2.280	-2.269	-2.270	-3.005	-3.003	-3.000	-2.981	-2.988

Note: Critical values are obtained based on 10000 simulations.

5 Average Lags

Table S5: Average Lag Order \bar{p} Chosen by AIC and SBC of the ADF Regressions for q_{ijct} With a Maximum lag of 12, When the Null Hypothesis is Rejected at 10% Significance Level in the Case with an Intercept and a Trend, for all 12 Countries

	$\bar{p}(\text{AIC})$	$\bar{p}(\text{SBC})$
CPI	4.33	1.17
Subcategories		
Bread	3.38	0.92
Meat	2.56	0.77
Dairy	2.23	0.74
Fruits	8.85	2.24
Tobacco	1.98	0.38
Alcohol	2.86	1.02
Clothing	8.77	6.73
Footwear	7.21	4.94
Rents	5.18	2.38
Fuel	2.73	0.65
Furnit.	5.15	2.67
Dom. Appl.	5.55	2.87
Vehicles	3.03	0.52
Pub. Transp	4.08	0.68
Comm.	2.15	0.62
Sound	4.42	0.64
Leisure	3.79	0.59
Books	2.11	0.48
Hotels	6.59	0.97
Average	4.35	1.62

Notes: This table is a supplement for Table 3 of the paper. The ‘Average’ located at the bottom row of the table is a simple average of the average lag orders chosen by AICs and SBCs for all the commodities except CPI.

6 Estimates of the Commodity Factor Models

Country	$\hat{\alpha}_{ic}$	$\hat{\gamma}_{ic1}$	$\hat{\gamma}_{ic2}$									
	Bread			Alcohol			Furnit.			Sound		
BE	1.43	-0.39	1.28	1.04	0.54	0.47	0.38	0.36	0.87	1.01	0.66	0.41
DE	-1.45	-0.26	1.10	-2.00	-0.17	1.18	-2.27	0.50	0.61	-2.02	0.61	0.46
DK	-0.79	-0.28	1.34	-0.21	0.50	0.35	-1.42	0.42	0.84	-0.87	0.56	0.56
ES	0.96	-0.34	1.76	1.69	2.23	-1.09	1.04	0.75	0.61	2.35	0.70	0.28
IT	5.02	1.73	-0.89	4.91	1.22	-0.34	3.91	1.21	0.00	4.79	1.03	-0.04
FR	-0.22	-0.13	0.94	-0.59	0.29	0.64	-1.37	0.45	0.74	-0.70	0.47	0.52
GR	2.85	-1.50	2.20	2.34	1.25	-0.42	1.79	0.46	0.58	1.95	0.68	0.35
NL	-0.89	-0.22	0.90	-1.55	0.19	0.71	-1.97	0.32	0.72	-1.15	0.58	0.23
PT	1.49	-1.16	2.42	1.47	1.77	-0.56	0.90	1.34	0.05	.	.	.
FI	-1.15	3.77	-2.77	-0.84	1.40	-0.49	.	.	.	-1.80	1.14	0.17
UK	-1.76	0.49	0.00	-2.53	0.50	0.27	-2.43	0.11	0.60	-1.79	0.71	-0.15
	Meat			Clothing			Dom. Appl.			Leisure		
BE	1.37	-0.20	1.09	0.57	1.22	-0.06	0.23	0.17	1.12	1.08	0.85	0.15
DE	-1.41	-0.23	1.04	-1.98	0.23	0.79	-2.42	0.28	0.87	-1.69	1.27	-0.35
DK	-0.26	0.20	0.67	-0.58	1.52	-0.54	-1.23	0.12	1.07	-0.80	0.31	0.73
ES	1.55	1.14	0.07	1.51	1.14	0.07	1.14	-0.16	1.49	1.75	3.12	-1.96
IT	4.87	0.74	0.14	4.89	1.39	-0.51	4.04	1.44	-0.28	3.77	-0.27	1.52
FR	-0.41	-0.25	1.10	-0.62	1.13	-0.19	-1.09	0.02	1.07	-0.38	2.22	-1.35
GR	3.60	-1.14	1.54	2.43	1.79	-0.97	2.29	0.12	0.73	2.78	5.90	-5.11
NL	-1.12	-0.46	1.20	-0.58	0.32	0.26	-2.11	0.12	0.96	-1.18	1.80	-1.00
PT	2.98	0.57	0.15	0.68	1.18	0.30	1.07	0.25	1.13	2.99	5.18	-4.39
FI	-2.47	3.41	-1.90	0.47	2.66	-2.17	.	.	.	-0.90	1.09	-0.17
UK	-1.41	-0.26	0.62	-2.67	-1.66	2.46	-2.30	-0.26	0.92	-2.06	5.46	-4.78
	Dairy			Footwear			Vehicles			Books		
BE	1.21	0.57	0.38	0.93	1.17	-0.13	1.33	0.92	-0.01	0.45	-0.49	1.68
DE	-1.58	-0.02	0.89	-2.02	0.91	0.12	-2.13	1.46	-0.39	-1.94	-0.15	1.13
DK	-0.64	1.43	-0.43	-0.63	1.17	-0.18	-0.80	0.70	0.36	-1.16	0.27	0.89
ES	1.51	1.09	0.13	1.70	1.35	-0.20	1.86	2.24	-1.15	1.75	0.32	0.79
IT	4.45	1.62	-0.58	4.76	1.54	-0.61	5.02	0.73	0.11	5.19	2.31	-1.50
FR	-0.42	0.34	0.52	-0.60	1.31	-0.38	-0.65	1.04	-0.09	-0.47	-0.33	1.19
GR	2.83	-2.37	3.03	2.96	2.08	-1.43	3.77	-2.39	2.77	0.96	0.72	0.61
NL	-1.44	0.01	0.84	-0.97	0.11	0.59	-1.84	1.20	-0.20	-1.36	-0.29	1.11
PT	2.37	1.67	-0.74	1.30	0.54	0.73	1.98	3.59	-2.58	0.51	-1.08	2.64
FI	-1.55	5.27	-4.07	-0.47	1.41	-0.62	-1.22	0.72	0.32	-0.30	4.06	-3.27
UK	-2.49	-0.55	1.28	-3.16	-1.45	2.41	-2.23	0.70	-0.05	-2.49	-0.21	0.94
	Fruits			Rents			Pub. Transp			Hotels		
BE	1.91	1.01	-0.29	1.92	2.02	-1.19	1.38	0.09	0.81	0.99	1.03	0.00
DE	-2.01	1.15	-0.13	-1.49	1.31	-0.37	-1.52	0.10	0.76	-1.38	1.44	-0.64
DK	-0.08	0.82	-0.02	-0.21	2.16	-1.16	-0.81	0.38	0.69	-0.14	1.56	-0.74
ES	1.35	0.79	0.49	3.23	1.76	-1.03	1.50	0.56	0.66	1.40	0.59	0.67
IT	4.55	1.03	-0.02	4.63	0.58	0.45	4.49	2.18	-1.18	4.51	1.35	-0.33
FR	-0.25	0.75	0.08	-0.29	1.86	-0.92	-0.29	0.06	0.77	-0.64	1.13	-0.20
GR	2.94	1.04	-0.40	3.94	4.69	-4.11	2.90	-0.56	1.23	2.59	0.24	0.54
NL	-1.12	0.97	-0.21	-0.95	2.16	-1.32	-1.29	0.05	0.78	-1.11	1.44	-0.70
PT	2.02	1.00	0.01	.	.	.	2.34	0.58	0.40	1.91	0.02	1.06
FI	-1.55	2.30	-1.20	-0.95	-0.55	1.50	-1.29	2.36	-1.31	-1.42	2.72	-1.64
UK	-2.52	1.40	-0.66	-2.67	1.54	-0.62	-2.22	0.96	-0.32	-2.20	1.56	-0.92

The table is continued on the next page.

Estimates of the Commodity Factor Models, Continued

Country	$\hat{\alpha}_{ic}$	$\hat{\gamma}_{ic1}$	$\hat{\gamma}_{ic2}$	$\hat{\alpha}_{ic}$	$\hat{\gamma}_{ic1}$	$\hat{\gamma}_{ic2}$	$\hat{\alpha}_{ic}$	$\hat{\gamma}_{ic1}$	$\hat{\gamma}_{ic2}$
	Tobacco			Fuel			Comm.		
BE	1.88	0.42	0.32	1.36	1.38	-0.49	0.90	1.24	-0.19
DE	0.25	0.03	0.24	-2.19	1.22	-0.15	-2.43	1.33	-0.17
DK	1.43	0.00	0.32	-1.57	0.52	0.77	-1.13	0.72	0.46
ES	3.47	0.28	0.27	1.69	1.36	-0.23	1.44	1.46	-0.24
IT	4.88	0.68	0.22	4.33	0.34	0.73	4.96	0.88	-0.01
FR	0.97	0.62	-0.24	-1.03	0.96	0.11	-0.34	1.36	-0.51
GR	2.90	0.58	0.08	1.26	2.31	-1.15	1.18	1.58	-0.41
NL	-0.01	0.34	0.07	-1.42	0.98	-0.12	-2.25	1.29	-0.15
PT	2.82	0.59	0.19	1.56	0.61	0.60	2.42	1.81	-0.90
FI	-0.51	-0.40	1.24	.	.	.	1.30	0.00	0.24
UK	-0.66	0.16	-0.02	-3.33	1.15	-0.14	-3.19	1.76	-0.78

Notes: This table reports the OLS estimates for the regressions $q_{ict} = \hat{\alpha}_{ic} + \hat{\gamma}_{ic1}\bar{q}_{ct} + \hat{\gamma}_{ic2}\bar{q}_t + \hat{\varepsilon}_{ict}$, $i = 0, 1, \dots, N_c$; $c = 1, 2, \dots, C$, $t = 1, 2, \dots, T_{ic}$, where $\bar{q}_{ct} = (N_c + 1)^{-1} \sum_{i=0}^{N_c} q_{ict}$ and $\bar{q}_t = C^{-1} \sum_{c=1}^C \bar{q}_{ct}$, for each combination of (i, c) .

7 Countries With Similar Factor Loadings

Table S7: Countries Whose Differences of Factor Loadings are ± 0.1 From Those of a Base Country

Category	Base country	Countries with ± 0.1 Difference
Bread	FR	NL
	NL	FR
Meat	BE	DE FR
	DE	BE FR
	FR	BE DE
Dairy	DE	NL
	NL	DE
Fruits	BE	NL
	IT	PT
	NL	BE
	PT	IT
Tobacco	DE	DK
	DK	DE
	IT	PT
	PT	IT
Alcohol	DK	UK
	IT	GR
	FR	NL
	GR	IT
	NL	FR
	UK	DK
Footwear	BE	DK
	DK	BE
Fuel	DE	UK
	UK	DE
Furnit.	BE	DK
	DE	GR
	DK	BE
	GR	DE
Dom. Appl.	BE	DK PT
	DK	BE
	PT	BE
Vehicles	DK	FI
	FI	DK
Pub. Transp	BE	DE FR NL
	DE	BE FR NL
	FR	BE DE NL
	NL	BE DE FR
Comm.	BE	DE NL
	DE	BE NL
	NL	BE DE
Sound	BE	DE GR
	DE	BE
	DK	FR
	ES	GR
	FR	DK
	GR	BE ES
Books	FR	NL
	NL	FR
Hotels	DE	NL
	NL	DE

Notes: The third column of the table reports the countries such that $|\hat{\gamma}_{ic1} - \hat{\gamma}_{jc1}|$ and $|\hat{\gamma}_{ic2} - \hat{\gamma}_{jc2}|$ are less than 0.1, where i is the base country and $j \neq i$ for each commodity c , $\hat{\gamma}_{ic1}$ and $\hat{\gamma}_{ic2}$ are factor loadings estimated from the regression $q_{ict} = \hat{\alpha}_{ic} + \hat{\gamma}_{ic1}\bar{q}_{ct} + \hat{\gamma}_{ic2}\bar{q}_t + \hat{\varepsilon}_{ict}$, $\bar{q}_{ct} = (N_c + 1)^{-1} \sum_{i=0}^{N_c} q_{ict}$ and $\bar{q}_t = C^{-1} \sum_{c=1}^C \bar{q}_{ct}$, for each combination of (i, c) .