

PAIRWISE TESTS OF PURCHASING POWER
PARITY USING AGGREGATE AND
DISAGGREGATE PRICE MEASURES

M. HASHEM PESARAN
RON P. SMITH
TAKASHI YAMAGATA
LIUDMYLA HVOZDYK

CESIFO WORKING PAPER NO. 1704
CATEGORY 10: EMPIRICAL AND THEORETICAL METHODS
APRIL 2006

An electronic version of the paper may be downloaded

- *from the SSRN website:* www.SSRN.com
- *from the RePEc website:* www.RePEc.org
- *from the CESifo website:* www.CESifo-group.de

PAIRWISE TESTS OF PURCHASING POWER PARITY USING AGGREGATE AND DISAGGREGATE PRICE MEASURES

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. Given data on $N + 1$ countries, $i, j = 0, 1, 2, \dots, N$, the standard procedure is to apply unit root or stationarity tests to N relative prices against a base country, 0, e.g. the US. The evidence is that such tests are sensitive to the choice of base country. In addition, the analysis is subject to a high degree of cross section dependence which is difficult to deal with particularly when N is large. In this paper 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 possible $N(N + 1)/2$ pairs of real exchange rate pairs between the $N + 1$ countries and estimate the proportion of the pairs that are stationary, for the aggregates and each of the 19 commodity groups. This approach is invariant to base country effects and the proportion that are non-stationary can be consistently estimated even if there is cross-sectional dependence. To deal with small sample problems and residual cross section dependence, 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 Code: C23, F31, F41.

Keywords: purchasing power parity, panel data, pairwise approach, cross section dependence.

M. Hashem Pesaran
Cambridge University
Department of Economics
Sidgwick Avenue
UK – Cambridge, CB3 9DD
Mhp1@cam.ac.uk

Ron P. Smith
Birkbeck College, London
School of Economics
Malet Street
UK – London WC1E 7HX
rsmith@econ.bbk.ac.uk

Takashi Yamagata
Cambridge University
takashi.yamagata@econ.cam.ac.uk

Liudmyla Hvozdyk
Cambridge University

March 30, 2006

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 $\lim_{T \rightarrow \infty} \Pr(ADF_{ijt}(p) < K_{T,p,\alpha} | H_N) = \alpha$. 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 $\{\hat{e}_{ct}\}_{t=1}^T$, 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 $\bar{q}_{ct}^{(r)} = \bar{q}_{c,t-1}^{(r)} + \sum_{\ell=1}^{s_c} \hat{\xi}_{c\ell} \Delta \bar{q}_{c,t-\ell}^{(r)} + e_{ct}^{(r)}$, $t = 1, 2, \dots, T$; $c = 1, 2, \dots, C$, $r = 1, 2, \dots, R$, $R = 2000$.

CESifo Working Paper Series

(for full list see www.cesifo-group.de)

- 1641 Stefan Bach, Giacomo Corneo and Viktor Steiner, Top Incomes and Top Taxes in Germany, January 2006
- 1642 Johann K. Brunner and Susanne Pech, Optimum Taxation of Life Annuities, January 2006
- 1643 Naércio Aquino Menezes Filho, Marc-Andreas Muendler and Garey Ramey, The Structure of Worker Compensation in Brazil, with a Comparison to France and the United States, January 2006
- 1644 Konstantinos Angelopoulos, Apostolis Philippopoulos and Vangelis Vassilatos, Rent-Seeking Competition from State Coffers: A Calibrated DSGE Model of the Euro Area, January 2006
- 1645 Burkhard Heer and Bernd Suessmuth, The Savings-Inflation Puzzle, January 2006
- 1646 J. Stephen Ferris, Soo-Bin Park and Stanley L. Winer, Political Competition and Convergence to Fundamentals: With Application to the Political Business Cycle and the Size of Government, January 2006
- 1647 Yu-Fu Chen, Michael Funke and Kadri Männasoo, Extracting Leading Indicators of Bank Fragility from Market Prices – Estonia Focus, January 2006
- 1648 Panu Poutvaara, On Human Capital Formation with Exit Options: Comment and New Results, January 2006
- 1649 Anders Forslund, Nils Gottfries and Andreas Westermark, Real and Nominal Wage Adjustment in Open Economies, January 2006
- 1650 M. Hashem Pesaran, Davide Pettenuzzo and Allan G. Timmermann, Learning, Structural Instability and Present Value Calculations, January 2006
- 1651 Markku Lanne and Helmut Luetkepohl, Structural Vector Autoregressions with Nonnormal Residuals, January 2006
- 1652 Helge Berger, Jakob de Haan and Jan-Egbert Sturm, Does Money Matter in the ECB Strategy? New Evidence Based on ECB Communication, January 2006
- 1653 Axel Dreher and Friedrich Schneider, Corruption and the Shadow Economy: An Empirical Analysis, January 2006
- 1654 Stefan Brandauer and Florian Englmaier, A Model of Strategic Delegation in Contests between Groups, January 2006

- 1655 Jan Zápál and Ondřej Schneider, What are their Words Worth? Political Plans and Economic Pains of Fiscal Consolidations in New EU Member States, January 2006
- 1656 Thiess Buettner, Sebastian Hauptmeier and Robert Schwager, Efficient Revenue Sharing and Upper Level Governments: Theory and Application to Germany, January 2006
- 1657 Daniel Haile, Abdolkarim Sadrieh and Harrie A. A. Verbon, Cross-Racial Envy and Underinvestment in South Africa, February 2006
- 1658 Frode Meland and Odd Rune Straume, Outsourcing in Contests, February 2006
- 1659 M. Hashem Pesaran and Ron Smith, Macroeconometric Modelling with a Global Perspective, February 2006
- 1660 Alexander F. Wagner and Friedrich Schneider, Satisfaction with Democracy and the Environment in Western Europe – a Panel Analysis, February 2006
- 1661 Ben J. Heijdra and Jenny E. Ligthart, Fiscal Policy, Monopolistic Competition, and Finite Lives, February 2006
- 1662 Ludger Woessmann, Public-Private Partnership and Schooling Outcomes across Countries, February 2006
- 1663 Topi Miettinen and Panu Poutvaara, Political Parties and Network Formation, February 2006
- 1664 Alessandro Cigno and Annalisa Luporini, Optimal Policy Towards Families with Different Amounts of Social Capital, in the Presence of Asymmetric Information and Stochastic Fertility, February 2006
- 1665 Samuel Muehleemann and Stefan C. Wolter, Regional Effects on Employer Provided Training: Evidence from Apprenticeship Training in Switzerland, February 2006
- 1666 Laszlo Goerke, Bureaucratic Corruption and Profit Tax Evasion, February 2006
- 1667 Ivo J. M. Arnold and Jan J. G. Lemmen, Inflation Expectations and Inflation Uncertainty in the Eurozone: Evidence from Survey Data, February 2006
- 1668 Hans Gersbach and Hans Haller, Voice and Bargaining Power, February 2006
- 1669 Françoise Forges and Frédéric Koessler, Long Persuasion Games, February 2006
- 1670 Florian Englmaier and Markus Reisinger, Information, Coordination, and the Industrialization of Countries, February 2006
- 1671 Hendrik Hakenes and Andreas Irmen, Something out of Nothing? Neoclassical Growth and the ‘Trivial’ Steady State, February 2006

- 1672 Torsten Persson and Guido Tabellini, Democracy and Development: The Devil in the Details, February 2006
- 1673 Michael Rauber and Heinrich W. Ursprung, Evaluation of Researchers: A Life Cycle Analysis of German Academic Economists, February 2006
- 1674 Ernesto Reuben and Frans van Winden, Reciprocity and Emotions when Reciprocators Know each other, February 2006
- 1675 Assar Lindbeck and Mats Persson, A Model of Income Insurance and Social Norms, February 2006
- 1676 Horst Raff, Michael Ryan and Frank Staehler, Asset Ownership and Foreign-Market Entry, February 2006
- 1677 Miguel Portela, Rob Alessie and Coen Teulings, Measurement Error in Education and Growth Regressions, February 2006
- 1678 Andreas Haufler, Alexander Klemm and Guttorm Schjelderup, Globalisation and the Mix of Wage and Profit Taxes, February 2006
- 1679 Kurt R. Brekke and Lars Sørsgard, Public versus Private Health Care in a National Health Service, March 2006
- 1680 Dominik Grafenhofer, Christian Jaag, Christian Keuschnigg and Mirela Keuschnigg, Probabilistic Aging, March 2006
- 1681 Wladimir Raymond, Pierre Mohnen, Franz Palm and Sybrand Schim van der Loeff, Persistence of Innovation in Dutch Manufacturing: Is it Spurious?, March 2006
- 1682 Andrea Colciago, V. Anton Muscatelli, Tiziano Ropele and Patrizio Tirelli, The Role of Fiscal Policy in a Monetary Union: Are National Automatic Stabilizers Effective?, March 2006
- 1683 Mario Jametti and Thomas von Ungern-Sternberg, Risk Selection in Natural Disaster Insurance – the Case of France, March 2006
- 1684 Ken Sennewald and Klaus Waelde, “Itô’s Lemma“ and the Bellman Equation for Poisson Processes: An Applied View, March 2006
- 1685 Ernesto Reuben and Frans van Winden, Negative Reciprocity and the Interaction of Emotions and Fairness Norms, March 2006
- 1686 Françoise Forges, The Ex Ante Incentive Compatible Core in Exchange Economies with and without Indivisibilities, March 2006
- 1687 Assar Lindbeck, Mårten Palme and Mats Persson, Job Security and Work Absence: Evidence from a Natural Experiment, March 2006

- 1688 Sebastian Buhai and Coen Teulings, Tenure Profiles and Efficient Separation in a Stochastic Productivity Model, March 2006
- 1689 Gebhard Kirchgaessner and Silika Prohl, Sustainability of Swiss Fiscal Policy, March 2006
- 1690 A. Lans Bovenberg and Peter Birch Sørensen, Optimal Taxation and Social Insurance in a Lifetime Perspective, March 2006
- 1691 Moritz Schularick and Thomas M. Steger, Does Financial Integration Spur Economic Growth? New Evidence from the First Era of Financial Globalization, March 2006
- 1692 Burkhard Heer and Alfred Maussner, Business Cycle Dynamics of a New Keynesian Overlapping Generations Model with Progressive Income Taxation, March 2006
- 1693 Jarko Fidrmuc and Iikka Korhonen, Meta-Analysis of the Business Cycle Correlation between the Euro Area and the CEECs, March 2006
- 1694 Steffen Henzel and Timo Wollmershaeuser, The New Keynesian Phillips Curve and the Role of Expectations: Evidence from the Ifo World Economic Survey, March 2006
- 1695 Yin-Wong Cheung, An Empirical Model of Daily Highs and Lows, March 2006
- 1696 Scott Alan Carson, African-American and White Living Standards in the 19th Century American South: A Biological Comparison, March 2006
- 1697 Helge Berger, Optimal Central Bank Design: Benchmarks for the ECB, March 2006
- 1698 Vjollca Sadiraj, Jan Tuinstra and Frans van Winden, On the Size of the Winning Set in the Presence of Interest Groups, April 2006
- 1699 Martin Gassebner, Michael Lamla and Jan-Egbert Sturm, Economic, Demographic and Political Determinants of Pollution Reassessed: A Sensitivity Analysis, April 2006
- 1700 Louis N. Christofides and Amy Chen Peng, Major Provisions of Labour Contracts and their Theoretical Coherence, April 2006
- 1701 Christian Groth, Karl-Josef Koch and Thomas M. Steger, Rethinking the Concept of Long-Run Economic Growth, April 2006
- 1702 Dirk Schindler and Guttorm Schjelderup, Company Tax Reform in Europe and its Effect on Collusive Behavior, April 2006
- 1703 Françoise Forges and Enrico Minelli, Afriat's Theorem for General Budget Sets, April 2006
- 1704 M. Hashem Pesaran, Ron P. Smith, Takashi Yamagata and Liudmyla Hvozdyk, Pairwise Tests of Purchasing Power Parity Using Aggregate and Disaggregate Price Measures, April 2006